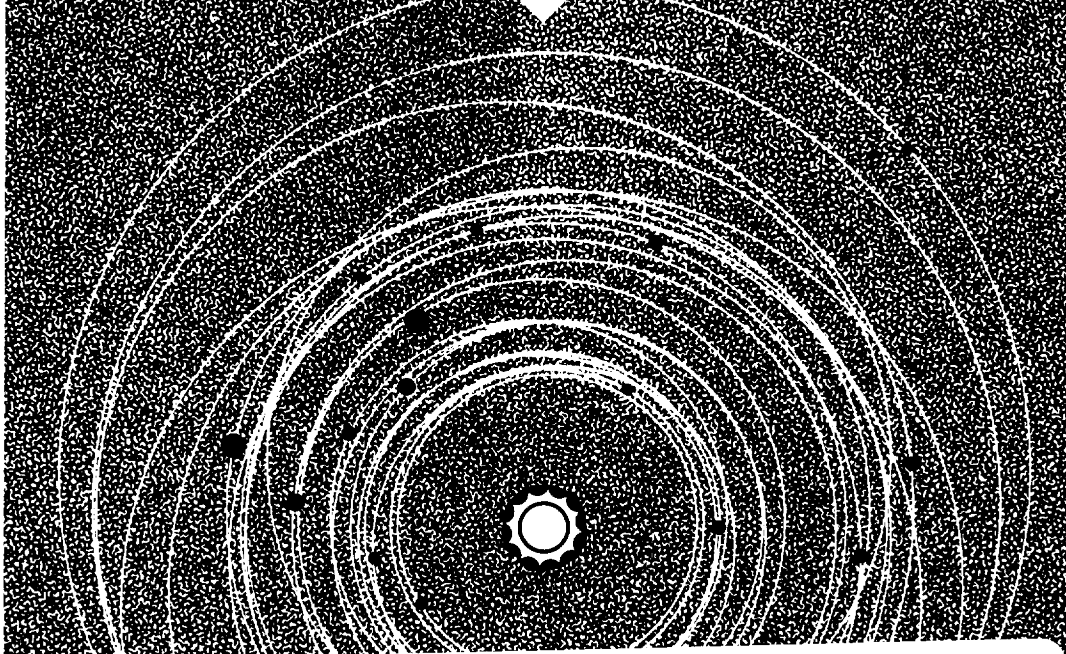
NASA-CR-193561

Strategy for the Detection and Study of Other Planetary Systems and Extrasolar Planetary Materials: 1990-2000



National Research Council

Strategy for the Detection and Study of Other Planetary Systems and Extrasolar Planetary Materials: 1990-2000

Committee on Planetary and Lunar Exploration
Space Studies Board
Commission on Physical Sciences, Mathematics, and Applications
National Research Council

NATIONAL ACADEMY PRESS
Washington D.C. 1990

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

This study was supported by Contract NASW-3482 between the National Academy of Sciences and the National Aeronautics and Space Administration.

Copies of this summary, as well as copies of the full report, are available from:

Space Studies Board
National Research Council
2101 Constitution Avenue, NW
Washington, DC 20418

Additional copies of the full report are available for sale from:

National Academy Press
2101 Constitution Avenue, NW
Washington, DC 20418

Printed in the United States of America

COMMITTEE ON PLANETARY AND LUNAR EXPLORATION

Larry W. Esposito, University of Colorado, *Chairman*
Robert O. Pepin, University of Minnesota (Chairman, 7/85-6/88)

Peter H. Bodenheimer,* University of California at Santa Cruz
Alan P. Boss, Carnegie Institution of Washington
Robert A. Brown,* Space Telescope Science Institute
Donald E. Brownlee,* University of Washington
Andrew F. Cheng, Johns Hopkins University
Anita L. Cochran, University of Texas at Austin
Paul D. Feldman,* Johns Hopkins University
Robert D. Gehrz,* University of Minnesota
William K. Hartmann,* Planetary Science Institute
Larry A. Haskin,* Washington University
Bruce M. Jakosky,* University of Colorado
Lucy-Ann McFadden, University of California at San Diego
Christopher P. McKay, NASA Ames Research Center
William B. McKinnon,* Washington University
H. Warren Moos,* Johns Hopkins University
Duane O. Muhleman, California Institute of Technology
Philip D. Nicholson,* Cornell University
Norman R. Pace, Indiana University
Carle M. Pieters,* Brown University
Graham Ryder, Lunar and Planetary Institute
Paul D. Spudis, Lunar and Planetary Institute
Peter H. Stone, Massachusetts Institute of Technology
David W. Strangway,* University of British Columbia
G. Leonard Tyler,* Stanford University
Fred L. Whipple,* Center for Astrophysics
Yuk L. Yung,* California Institute of Technology
Richard W. Zurek, Jet Propulsion Laboratory

Staff

Paul F. Uhlir, Senior Program Officer
Anne L. Pond, Secretary
Altoria L. Bell, Secretary

---

*Former committee members who participated in writing or reviewing this report.

SPACE STUDIES BOARD

COMMISSION ON PHYSICAL SCIENCES, MATHEMATICS, AND APPLICATIONS*

Norman Hackerman, Robert A. Welch Foundation, *Chairman*
Peter J. Bickel, University of California, Berkeley
George F. Carrier, Harvard University
Herbert D. Doan, The Dow Chemical Company (retired)
Dean E. Eastman, IBM T.J. Watson Research Center
Marye Anne Fox, University of Texas
Phillip A. Griffiths, Duke University
Neal F. Lane, Rice University
Robert W. Lucky, AT & T Bell Laboratories
Christopher F. McKee, University of California at Berkeley
Richard S. Nicholson, American Association for the Advancement of Science
Jeremiah P. Ostriker, Princeton University Observatory
Alan Schriesheim, Argonne National Laboratory
Roy F. Schwitters, Superconducting Super Collider Laboratory
Kenneth G. Wilson, Ohio State University

Norman Metzger, Executive Director

*The project that is the subject of this report was initiated under the predecessor group of the Commission on Physical Sciences, Mathematics, and Applications, which was the Commission on Physical Sciences, Mathematics, and Resources, whose members are listed in Appendix B.

Contents

Executive Summary

The recent decades of planetary exploration by spacecraft have centrally influenced perception and understanding of our solar system. But fundamental questions remain, questions of origin and uniqueness that cannot be answered with confidence as long as we are limited by access to just this one example of a solar system. This report addresses a new opportunity in the planetary sciences—to extend our exploration outward to discover and study planetary systems that may have formed or are forming around other stars.

In this report, the Committee on Planetary and Lunar Exploration (COMPLEX) concludes that a coordinated program of astronomical observation, laboratory research, and theoretical development aimed toward systematically exploring the solar neighborhood for planets and planetary systems in various stages of evolution around other stars, and understanding the dynamics and origin of whatever may be found, would be a technologically feasible, scientifically exciting, and potentially richly rewarding extension of the study of bodies within the solar system. It has, in fact, already begun. Discoveries of orbiting bodies with masses far below those of the smallest previously known stars have been announced even though these bodies lie at the edge of detectability and await confirmation. Thus, within just the past few years, a provocative and challenging new opportunity has appeared in the planetary sciences—the opportunity to readdress fundamental scientific and intellectual questions of origin by moving beyond the limited knowledge afforded by the single example of our own planetary system to a more general understanding based on the existence (or, conceivably, nonexistence) and character of other evolved systems.

In the opinion of the committee, the scientific purview of such a program necessarily includes the antecedents of evolved planetary systems as well. The program, therefore, should encourage and support relevant observational, laboratory, and theoretical investigations of precursor matter and of young dusty stellar environments that are likely sites for planet formation. Studies of this nature involve a variety of observational techniques within several different areas of astronomical and astrophysical research, and this component of the program is thus intrinsically multidisciplinary. So, too, is the effort to detect and characterize evolved planets, since the instrumental capabilities needed for such observational searches are directly applicable to current endeavors and problems in galactic astronomy, astrophysics, and cosmology. To note one example, an astrometric telescope with the sensitivity recommended in this report would have a substantial impact on determining stellar distances by direct parallax measurements; it could, for instance, yield accurate distances to at least 20 "standard candle" Cepheid variables and thus precisely calibrate the distance scale in the local group of galaxies and improve that of the universe as a whole.

The committee concludes, from consideration of the scientific and engineering issues involved and from assessment of existing and developing detection techniques, that the relevant technology is on the brink of major advances; that it will be technically feasible, on a one- to perhaps three-decade time scale, to make a first assessment of whether extrasolar planets in the Jupiter- to Uranus-mass range are common or very rare; and that resolution of this question is of immense scientific importance. Initiation and support of a formal program designed to answer this question, and to probe all stages of planetary evolution illuminated by the astronomical record, are therefore, in the committee's view, worthy of major scientific effort at this time.

In Chapters 5, 6, and 8 of this report, COMPLEX has set out recommendations for the overall strategic considerations, including scientific objectives and requirements for the specific measurements that are needed to achieve them and that in the committee's opinion should guide the structure, content, conduct, and time scales of such a program in the near term. *In summary, the recommendations to the NASA Office of Space Science and Applications (OSSA) are to:*

- *Initiate, and maintain for at least a decade, systematic observational planet searches that encompass the widest feasible domain of the planetary mass versus semimajor axis discovery space (see Figure 5.2 in Chapter 5). Specifically,*

1. *Initiate an astrometric observational survey program designed to track the reflex motion of 100 or more stars in the solar neighborhood ($r \leq 10$ parsecs) with a design goal for relative astrometric accuracy of $\sigma = 10$ microarcsec, sufficient in a search of adequate duration to detect and track Uranus-mass planets in a solarlike system.*

2. *Obtain and interpret a record of Doppler shifts in stellar spectral features due to reflex motion, at or above the current measurement accuracy of $\sigma \cong 10\ m\ s^{-1}$ for the velocity of the orbital reflex, in a survey of the duration and extent specified for the astrometric survey.*

3. *Until such systematic searches are mounted, maintain ongoing ground-based searches at their present best accuracies, and investigate and implement improvement of these accuracies if technically and financially feasible.*

- *Augment current observational studies of young stellar systems, and of the physical properties of circumstellar-interstellar dust systems as precursors to and products of planetary systems, on a variety of spatial and spectral resolution scales. Survey a statistically meaningful number of stars of varied masses and types to detect such systems.*
- *Continue investigations of links between interstellar-circumstellar dust and isotopically "exotic" grains in solar system materials such as primitive meteorites, interplanetary dust, and comets. Important elements and objectives of this effort include collection and curation of rare interplanetary asteroidal-cometary dust particles (IDPs); laboratory identification and analysis of micron to submicron presolar dust grains preserved in these meteoritic and IDP materials; and laboratory simulation and theoretical studies of the astronomical dust cycle, including the formation and physical and chemical processing of interstellar grains in preplanetary and planet-forming environments.*
- *Improve the capability of theoretical models and computer experiments to make specific predictions regarding the observational properties of planetary systems at all stages of their evolution, and to further develop models to aid in the interpretation of existing data.*
- *Encourage the following multi-disciplinary activities between the responsible divisions at OSSA: participation of planetary scientists in the design and building of future observatories and facility instruments, and in the allocation of observing time at existing observational facilities; joint support for multi-disciplinary scientific initiatives; and joint development of instrumentation for extrasolar observation.*
- *Pursue long-range instrumental and strategic initiatives that are conceptually applicable and potentially valuable to the investigation of extrasolar planetary materials in later stages of reconnaissance or in subsequent phases of exploration and intensive study, but that at present are technologically or theoretically too undeveloped to be of immediate utility in implementing the short-term strategy proposed in this report.*

1

Introduction

The recent decades of planetary exploration by spacecraft have centrally influenced perception and understanding of our solar system. But fundamental questions remain, questions of origin and uniqueness that cannot be answered with confidence as long as we are limited by access to just this one example of a solar system. This report addresses a new opportunity in the planetary sciences—to extend our exploration outward to discover and study planetary systems that may have formed or are forming around other stars.

The intellectual desire to look beyond the confines of our own system for answers to these unknowns is of course not new, but what has emerged only recently is the chance to actually find them. The questions to be asked about other planetary systems have been posed and sharpened against a broad and growing base of local knowledge just when the technological ability to detect other systems is poised on the brink of actual discovery. Until the 1980s there was no real evidence for the existence of small condensed orbiting bodies, other than tiny stars, in other stellar systems. But in just the past few years a number of viable candidates have been found, some perhaps as small as Jupiter. We are truly at the edge of a scientific and philosophic revolution.

Investigation of the existence and nature of extrasolar planets, and of the matter from which they form and evolve, is unquestionably an endeavor of rich scientific interest and consequence to planetary science, astronomy, and astrophysics. It also involves the profound philosophic question of the relation of Earth and humanity to the rest of the universe. Histories of science—indeed histories of human intellectual development during and after the Renaissance—emphasize four major earlier revolutions, each of which displaced humanity ever further from a seemingly central and unique position of place and perception. Two of these were purely astronomical. First, the Copernican revolution overturned the Aristotelian "common sense" view that Earth occupies the center of the universe. Much later, in the 1920s, what might be called the "Shapley revolution," after one of the principal astronomers involved, showed that the solar system was not at the center of our galaxy, nor was the disk-shaped galaxy a unique system including all known stars and "spiral nebulae" as had been thought. Rather, the solar system was out toward the edge, in one arm of our spiral galaxy. And so-called spiral nebulae had been misidentified. They were not spiral nebulae at all, but distant galaxies, and our own galaxy was just one of many.

Between these two revolutions in our perception of cosmic place came two more, one biological and the other physical. Charles Darwin showed that we were not caretakers of creation, as had been imagined in earlier Western thought, but rather one of many interdependent, evolving species in a long history of species that had emerged, changed, and died. And with the advent of relativity and quantum mechanics from Albert Einstein, Max Planck, Niels Bohr, and Werner Heisenberg, notions of a privileged human observer, direct perception of physical law, and a classically deterministic universe disappeared forever.

Each of these revolutions had the effect of displacing humankind from its assumed anthropocentric position. In this context one dramatic question remains. Is it possible that Earth, the habitable conditions on Earth, and indeed the life that has evolved to fit those conditions constitute rare accidents? Modern theorists have proposed answers, but we will never really know, in a scientific sense, until we have surveyed a statistically valid sample of star systems with enough sensitivity to determine whether they have Jupiter-size, Uranus-size, and ultimately even Earth-size planetary bodies near them. Barring interception of signals from intelligent life elsewhere by radio listening searches, only such a survey can address this question.

Either a positive or negative answer would have profound consequences. The search might reveal that planetary bodies are common stellar companions, existing near 10 percent, 50 percent, or even 90 percent of all stars. Then we would find ourselves in the midst of a fundamental extension of earlier revolutions, moving us still further from an anthropocentric view in demonstrating that our home system of planets and perhaps life itself are not unique or special creations. Scientifically, such a result would put much theoretical work on a firmer footing by clarifying origins and allowing statistical studies. A negative result—no detections after a sensitive search among a large sample of stars—would be equally or perhaps even more profound, both intellectually and scientifically. It would certainly lead to basic reassessments in our thinking about solar system origin, the possibility of life elsewhere, and the general processes governing the evolution of stellar systems.

PREMISES OF THIS REPORT

In 1975, 1978, and 1980, the Committee on Planetary and Lunar Exploration (COMPLEX) of the Space Science Board (SSB, now the Space Studies Board) published three reports, which taken together encompass our entire planetary system. The first, "The Outer Planets," is included in the Space Science Board's *Report on Space Science 1975* (National Academy of Sciences, 1976). *Strategy for Exploration of the Inner Planets: 1977–1987* and *Strategy for the Exploration of Primitive Solar-System Bodies—Asteroids, Comets, and Meteoroids: 1980–1990* were published by the National Academy of Sciences in 1978 and 1980, respectively. The strategies for exploration of the outer solar system and of the inner planets were subsequently updated by two reports: *A Strategy for Exploration of the Outer Planets: 1986–1996* (National Academy Press, 1986) and *1990 Update to Strategy for Exploration of the Inner Planets* (National Academy Press, 1990). These five reports are the committee's principal advisory documents on scientific exploration of the solar system.

In its 1980 report, COMPLEX set forth general goals for investigation of primitive bodies within the solar system. The committee then looked briefly outward to a broader context in which exploration of our planetary system is exploration of just 1 of some 10^{11} stars in the galaxy. That short discussion emphasized the relationships that must exist between the history of our own solar system and the origin and evolution of other stellar systems. In highlighting the essential role that extrasolar observations will eventually play in our understanding of solar system history, this section of the 1980 report foreshadowed the current interest and activity in the detection and study of extrasolar planetary material, and it stands as a direct antecedent to the present report.

Remarkable observational advances during the past few years have pointed to the presence of solid matter at various stages of development around other stars. The whole-sky survey from the Infrared Astronomical Satellite (IRAS) provided compelling evidence for the presence in nearby star-forming regions of dusty matter surrounding a significant fraction of young ($t < 10$ million yr) solar-type, pre-main-sequence stars. The infrared spectra of such stars suggest that these materials occur in disklike circumstellar structures of dimensions comparable to the solar system and masses comparable to a "minimum mass" solar nebula. The IRAS also found spatially resolved dust shells (or rings) around several older, main-sequence stars, among them Vega, Beta Pictoris, and Fomalhaut. In the particular case of Beta Pictoris, a disklike structure has actually been detected. Extensive investigations of this system indicate that the dust grains are larger than typical interstellar grains, implying that they may have experienced growth processes similar to those that operated in the early solar system. The detection of star-orbiting bodies with masses significantly below those of the known lowest-mass stars is now technically feasible, by direct imaging, spectroscopy, or orbital reflex effects, and reports of the existence of such objects are appearing at an increasing rate. Continuing discoveries of circumstellar structures and objects relevant to this study can be expected to lead to rapid and fundamental changes in our understanding of this field and consequently to parallel modification of the specific search strategy proposed in this report. The committee expects, however, that the general approach and techniques discussed here will continue to be appropriate.

In light of recent observations and the potential for future discoveries, the SSB requested in 1985 that COMPLEX extend its consideration of exploration strategies to planetary systems outside the solar system. The committee was charged to assess both the significance of detection and study of extrasolar planetary material to our understanding of the origin and evolution of our own solar system, and the status and prospects for development of techniques that could be applied to investigation of objects far beyond direct access by spacecraft.

The base of information for the recommendations of this report was developed during 1984 to 1987 in COMPLEX deliberations combined with presentations by scientific and technical experts from outside the committee. Of particular importance were extensive briefings by participants in a series of workshops, sponsored by the National Aeronautics and Space Administration (NASA) from 1976 to 1984, on the scientific rationale for and feasibility of astrometric, Doppler spectrometric (radial velocity), photometric, and direct-imaging techniques for detection and study of extrasolar planetary material. After evaluating the information from these and other sources, COMPLEX has formulated the present report, which is in response to the charge from the Space Studies Board:

1. *Address the scientific rationale and goals* for investigations of other planetary systems and extrasolar planetary material in the context of current knowledge and the framing of scientific questions concerning the origin and evolution of planetary systems in general, and of our solar system in particular;
2. *Examine the status of theoretical understanding* of star and planet

formation, the state of relevant observations, the interplay between observational and theoretical studies, and the relevance to related areas of astronomical, astrophysical, and exobiological research of a program to detect and study other planetary systems;

3. *Assess the status of current research* in this and related areas, the evidence from classical and newly developed techniques, and the significance of recent discoveries;

4. *Consider and evaluate the observational capabilities*—existing, under development, and conceptual—needed for the measurements required to meet the science objectives; and

5. *Propose a strategy* for the investigation of other planetary systems and extrasolar planetary materials for the decade 1990 to 2000.

In developing this strategy and recommendations for its implementation, COMPLEX has adopted the position, also taken in its previous strategy reports, that long-term science objectives are most flexibly and enduringly defined in ways that separate strategic approaches from the specific means chosen by NASA to implement them. The committee wishes, however, to reiterate here a point emphasized in its 1978 report, that this or any other strategy of scientific exploration is meaningful only to the extent that it is actually carried out. Implementation requires operational planning, funding, and execution. Execution of the strategy depends on missions; in the present context this committee broadly defines "missions" to include development, deployment, and support of earth-based and earth-orbital instruments and facilities.

OVERVIEW AND PURVIEW

From direct observations, we see remarkable structural and chemical regularities within our solar system: low eccentricities and inclinations of planetary orbits, and coherent radial trends in composition. Similar regularities are displayed, in miniature, within the major satellite systems of the giant planets Jupiter and Saturn. These observations suggest that the coherent structure of at least this system was not an accident of random events.

In the hierarchy of the planetary sciences, the next level of inquiry asks if our system is one of a broad set with similar characteristics, or if formation of our planetary system around the Sun was a result of extraordinary processes. The implications of this question are intellectually profound and scientifically fundamental to our understanding of solar system genesis. It is difficult to address without data on the frequency of occurrence, structure, and dynamics of planetary systems around other stars. No matter what the answer, we will learn much about the origin and early evolution of the solar system from investigation of extrasolar planetary and preplanetary systems, whether it turns out that our system is a member of a common class and can be studied as such, or is rare or even unique. The imperative to understand the conditions and processes that culminated in the formation of our planetary system is the rationale for searching outward for answers.

In this context, investigation of extrasolar planetary materials represents a scientifically compelling initiative in the planetary sciences, a necessary outreach of a mature discipline. General understanding of the formation of planetary systems is evolving rapidly. It is strongly—and perhaps correctly—influenced by observed regularities in the "local example" and by evidence that rotating disks of gas and dust are present around pre-main-sequence stars. The evidence of rotating disks is consistent with concepts of precursor environments in which various physical processes, such as nonaxisymmetric gravitational effects, viscous dissipation, and magnetic fields, could have generated such regularities. Confirmation and observation of planets around other stars would allow more general characterization of the results of planet-forming events and would clearly lead to further understanding of the possible mechanisms involved.

The question of whether other planetary systems exist, as well as how they are ordered if they do exist, is difficult to constrain theoretically or by computer experiments. Observational searches are required at levels of capability and duration that can reveal, within scientifically interesting limits, not only the presence or absence of planets, but also the structure of systems that may be discovered. Therefore, COMPLEX approached this study with the view that, while simple detection of a planetary object is of great philosophic and public interest, only a continued program to characterize the properties of planetary systems in a statistical sense will allow us to address the most significant scientific questions.

In the following chapters of this report, COMPLEX sets out the basis of its recommendations for a program to detect and study other planetary systems and extrasolar planetary materials. (The shorter term "extrasolar planetary materials" is generally used throughout this report in place of the more cumbersome phrase "other planetary systems and extrasolar planetary materials," except where the context indicates otherwise.) The detection of evolved planetary materials requires new and renewed emphasis on observational capabilities. The intrinsic scientific importance and observational challenges of this endeavor have already catalyzed significant improvements in instruments and techniques. Initiation of a focused program of investigation at this time seems certain to lead to dramatic instrumental and analytic advances and to major new insights.

The committee proposes a decadal time frame for the strategy developed here. This choice was made not simply because it is congruent with the duration assigned to previous COMPLEX exploration strategies, but because the committee believes that this program will require a thorough reassessment in 10 years. Many circumstellar bodies should have periods of revolution that are less than a decade, and for them even relatively brief observational searches may be expected to provide unambiguous information. Others, such as Uranus and Neptune in our own system, have longer periods, and more extended observations will be necessary. For this reason and others, the committee has stressed the importance of the duration as well as the accuracy of the astronomical searches in developing measurement requirements. But the disadvantages of attempting to formulate a detailed and responsible strategy with open-ended duration in this rapidly evolving area, based as it would necessarily be on doubtful assumptions concerning future technological developments and potential breakthroughs in observational instrumentation and techniques, were in the committee's opinion compelling. The decadal time frame was therefore chosen to trigger a review of the progress and prospects of the program at a time when much of the recommended observational technology should be in place, and the data base applicable to various aspects of the investigation more firmly developed. COMPLEX emphasizes that there is no implication in this choice of strategic time frame that initiation, evolution, and termination of the research effort would all be accomplished within 10 years.

Many of the various research areas discussed in this report, in the context of a specific program for investigation of extrasolar planetary materials, are of necessity intrinsically interdisciplinary in nature and have broad observational and theoretical connections with other fields of scientific endeavor. Comprehensive investigation of precursor materials, for example, extends deeply and quantitatively into astronomical, astrophysical, and chemical studies of molecular clouds and of cloud evolution toward condensation at their cores. In this report, COMPLEX has focused its attention primarily on observation and study of post-collapse or post-condensation development, rather than on more general aspects of molecular cloud systems. Even so, essential elements of the investigation recommended here interweave in significant ways with this and other associated fields of scientific pursuit, and in the committee's opinion, new NASA-wide institutional structures involving the Solar System Exploration, Astrophysics, and eventually Life Sciences divisions will be required to fully implement these cross-disciplinary connections.

In this context, it is important to note specifically the purview of this study and of the exploration strategy derived from it. The committee includes considerations and recommendations pertaining to the following:

- Searches for extrasolar-system planets, preplanetary precursor systems, and other orbiting materials or objects of less than stellar mass;
- Determination of the characteristics of any planetary systems that may be discovered;

- Measurement of system elements in ways that facilitate comparisons with planets or dust in other systems;
- Characterization (spatial, spectral, and temporal) of young stellar systems that may be in the process of forming planetary systems; and
- Study of the physical and chemical nature of materials that primitive stellar systems may comprise.

The committee excludes or discusses only briefly and makes no detailed recommendations concerning the following areas of scientific or technological study:

- The general history of molecular clouds prior to condensation into protostellar systems;
- The capabilities of specific current and proposed astronomical facilities to probe the dynamics of molecular clouds, and the measurement capabilities required to enable such studies in the future;
- The cosmic history of the biogenic elements; and
- Programs focused principally on the search for extraterrestrial intelligence (SETI).

The following section of this introductory chapter defines the technical nomenclature adopted by the committee for the preparation of this report. Chapter 2 sets out the fundamental scientific goals of a program to detect and study extrasolar preplanetary materials, planets, and planetary systems. Chapter 3 summarizes current theoretical understanding of stellar and planetary formation, and Chapter 4 the present state of our observational knowledge of the precursors, processes, and products of possible planet-forming astronomical environments. Chapter 5 addresses methods and requirements for detecting and tracking planets, including currently feasible technological capabilities to mount and sustain credible searches for planets around neighboring stars. Specific measurement requirements are defined and developed, with emphasis on astrometric and Doppler spectroscopic instruments and techniques for detecting stellar reflex motion, which accounts for one component of the search strategy. Chapter 6 deals with ongoing and future physical studies of extrasolar materials that are possible precursors of evolving and evolved planetary systems, and their scientific and programmatic relevance to the strategy recommended in this report. Connections between this search for planetary and preplanetary objects and other fields of astronomical and astrophysical research, and the influence on some of these disciplines of the new observational capabilities intrinsic to planet-searching instruments, are discussed in Chapter 7. In Chapter 8 COMPLEX presents its recommended scientific objectives, measurement requirements, and overall decadal strategy for investigation of extrasolar planetary materials.

NOMENCLATURE

Planetary matter around other stars may be expected to span a vast range of mass and size. The search for such material is not limited to possible "Jupiters" or "Earths." Planetary materials can be thought of not only as any solid matter accreted to planetary size, but also, in earlier evolutionary stages, as dispersed dust grains condensing from the gas phase of a circumstellar accretion disk or infalling from precursor cloud environments. The following review of phenomena encountered at different masses, starting with small material, discusses each mass interval as it relates to the question of extrasolar planetary material. The terminology is that chosen by the committee to simplify the discussion, to allow framing of useful research questions, and at the same time to be as congruent as possible with existing literature. The nomenclature is summarized in Table 1.1 and is schematically illustrated in Figure 1.1.

According to current theory and observation, as a star forms it is surrounded by a cloud of gas and dust. One such cloud was the solar nebula, in which our own planetary system formed. Current observations and models indicate that interstellar dust grains were preserved in much of this cloud, and cooling produced condensation of even more grains. Such dust-rich clouds are inferred to exist around many solar-type pre-main-sequence stars, from observation of excess infrared radiation emitted by heated circumstellar grains. Particles of mass smaller than about 1 μg ($d < 10^{-2}$ cm) are conveniently called *dust*. An important research question is whether observed circumstellar dust is currently actively aggregating into planetary bodies, undergoing removal from a system without accretion, or being produced by erosion of already-formed, larger asteroidlike or cometlike bodies.

Theory suggests that gravitational settling of dust grains and aggregates toward the equatorial plane of a forming solar system, together with collisions, leads to the gravitational clumping or mechanical accretion (by mutual impact) of these grains into asteroid-sized objects. Ultimately, as is thought to have been the case in our solar system, accretion may produce planets at least as large as the terrestrial planets, and perhaps as large as the cores of the giant planets (around 10 to 20 $M_\oplus$). A physical basis for subdividing the nomenclature comes from noting that the strength of bodies smaller than a few hundred kilometers in diameter is typically greater than the internal gravitational stresses, so that they can maintain nonspherical shapes. Bodies larger than $\sim 10^{-2}$ cm but smaller than a few hundred kilometers are too big to be called dust, but they are smaller and usually more irregular in shape than are objects typically considered planets. Star-orbiting bodies in this mass range are called *subplanetary objects*. In our solar system, asteroids and comet nuclei are examples, but in this report

TABLE 1.1 Nomenclature

Object	Upper Mass Boundary	Boundary Rationale	Examples	Special Terms
Dust	10^{-6} g	Arbitrary	Meteoroidal and cometary dust, zodiacal light particles, interstellar grains	
Subplanetary object	2×10^{23} g M_{Vesta} $10^{-10}\ M_{\odot}$	Gravity dominates over strength of material in determining shape	Asteroids, comets, meteoroids, small satellites	*Planetesimal*
Planet	4×10^{30} g $2\ M_{Jupiter}$ $2 \times 10^{-3}\ M_{\odot}$	Electron degeneracy dominates over electrostatic effects in determining structure	Solar system planets and major satellites	*Planetesimal, protoplanet*
Substellar object	1.5×10^{32} g $80\ M_{Jupiter}$ $8 \times 10^{-2}\ M_{\odot}$	Hydrogen-burning nuclear reactions dominate energy generation	No *confirmed* examples as of 1990	*Brown dwarf*
Star	-	-	10^{11} examples in galaxy	*Protostar, pre-main-sequence star*

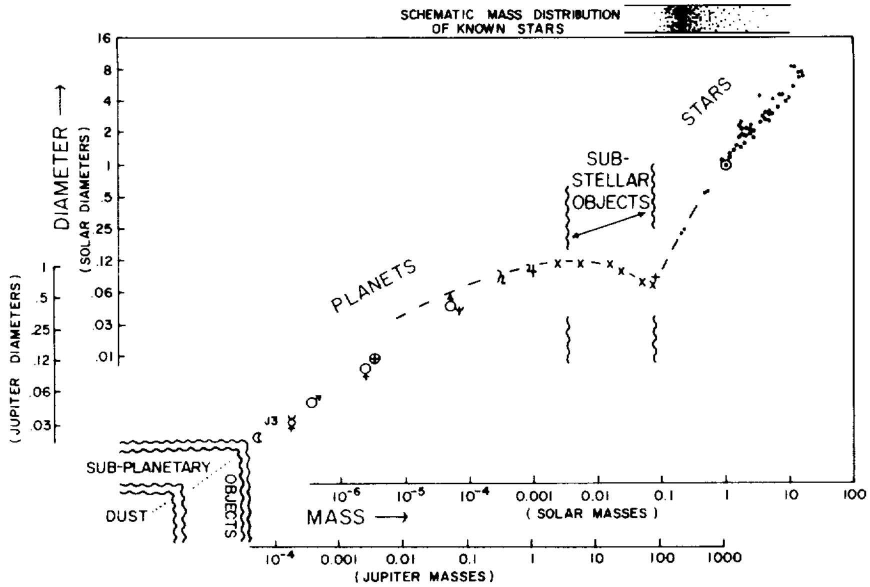


FIGURE 1.1 The various classes of target objects in the search for extrasolar planetary materials and systems, arranged schematically on a diagram of mass versus diameter. Dots in upper right give empirical data on main-sequence stars of solar composition (Popper [1980], *Annu. Rev. Astron. Astrophys.* 18, 115). Below the mass cutoff for nuclear reactions at about 80 $M_{Jupiter}$, theoretical models at minimum radius for solar composition are shown by plus and multiplication symbols (respectively from D'Antona [1987], *Astrophys. J.* 320, 633; and Burroughs, Hubbard, and Lunine [1989], *Astrophys. J.* 345, 939. The decline from a maximum near 2 $M_{Jupiter}$ to a minimum near 80 $M_{Jupiter}$ is due to compression as more mass is added. This is the regime of "substellar" objects in this report's nomenclature. Empirical data for the "planetary" regime are shown by planetary symbols; two large satellites are included. Two other regimes, of "subplanetary" objects and "dust," fall off scale to the lower left. Chart above the upper figure boundary denotes the observed mass distribution of stars, with a peak around 0.25 $M_\odot$ and a cutoff (observational) near 0.1 $M_\odot$.

the committee generally avoids applying these terms to other star systems to avoid connotations of composition or origin. Extrasolar subplanetary objects will not be individually observable in any direct way, although it is possible that their presence could be inferred. For example, it has been suggested that inner portions of the Beta Pictoris system that are seemingly empty of dust may be zones where dust has aggregated into larger bodies in the subplanetary-mass range; if so, the resulting configuration might be similar to our asteroid belt.

The term "planet" carries the connotation of a larger, relatively spherical body in a mass range at least as great as that from Mercury to Jupiter.

No universally accepted definition is available. Therefore for this discussion, a *planet* is defined as a star-orbiting object (of no presupposed, particular mode of origin) with a mass greater than that required for the development of a near-spherical shape (roughly the mass of the asteroid Vesta) but smaller than the upper limit in mass discussed below. Planets are thus generally larger than the asteroids of our own solar system, although Ceres would qualify in this nomenclature as a very small planet.

The upper limit on planet mass was chosen as follows. If sufficient mass is added to a body, normal crystalline or liquid matter becomes compressed and electron degeneracy becomes dominant. As seen in Figure 1.1, the radius of the body actually begins to decline with increasing mass at about this point. The precise position of turnover in the mass-radius diagram depends on composition and rotation and is thus somewhat ill-defined. However, it occurs near 2 $M_{Jupiter}$ for a range of pure hydrogen and solar-composition models. This transition, at about 2 $M_{Jupiter}$, is thus chosen as the physical basis for defining the upper limit for planets and the lower limit for substellar objects. This limit also corresponds well with common conceptions of what a planet is.

The boundary between substellar and stellar states is similarly (and, in this case, traditionally) defined by a physical transition. At about 80 $M_{Jupiter}$, hydrogen-burning nuclear reactions in the interior of the body can supply the entire radiated energy. Objects more massive than this are defined as stars. Bodies with a mass between 2 and 80 $M_{Jupiter}$ have sometimes been called "brown dwarfs," because they contract for cosmically long times and radiate in the infrared; however, the radiation is for the most part gravitational in origin, not nuclear. Traditionally, no lower mass limit has been indicated for the term brown dwarf, and considerable controversy has been engendered as to whether certain objects should be called brown dwarfs or planets. Because the term brown dwarf has been controversial as well as vague, it is not used in this report. Instead, a *substellar object* is defined as having a mass between 2 and 80 $M_{Jupiter}$, and a *star* as having a mass greater than 80 $M_{Jupiter}$.

The purpose of this terminology is not to create arbitrary categories, but rather to employ a language that simplifies discussion, is relatively congruent with common usage, and does not connote specific modes of origin. Definitions are based entirely on a mass spectrum, divided (except at the dust-subplanetary boundary) by physical phenomena. Subdivisions involving composition or other properties could be added, but Table 1.1 is adequate for the present discussion. The terms *substellar* and *subplanetary* are parallel and avoid the preconceptions associated with terms such as "brown dwarf," "asteroid," or "comet." Implications of origin are likewise avoided by defining an object as a planet solely on the basis of mass range.

Several other widely used special terms—*planetesimal*, *protoplanet*, *protostar*, and *pre-main-sequence star*—have not been directly included in this set of definitions, but are listed in Table 1.1 according to the masses these objects are usually considered to have. They appear in this report in contexts where their usage is conventional and their meanings clear. The term *planetesimal*, connoting young or preplanetary bodies with a potential for further aggregation, is used in a general way to designate objects in the subplanetary-mass or low-planetary-mass range, usually in reference to a system where there is a potential for collisional or gravitational interaction leading to evolution of the system as a whole. *Protoplanet* refers to an object of planetary mass at an early stage of development, a body with a substantial fraction of its final mass serving as a nucleus for accumulation of more material. Similarly, a *protostar* is an object of stellar mass in its earliest evolutionary stage, during which hydrodynamic accretion onto an equilibrium core occurs. A *pre-main-sequence star* refers to a later phase when the object is entirely in hydrostatic equilibrium, but has too low an internal temperature for generation of appreciable nuclear energy.

In general, important scientific considerations and questions, particularly as they relate to stars, substellar objects, and planets, are communicated readily and clearly with this terminology. One can frame the complicated question of whether planets form by different mechanisms, such as accretion, gravitational collapse, or capture, and the observational question of whether substellar objects or planets in any given system are in nearly coplanar and circular orbits. If a star has only one substellar or planetary companion, does that body orbit in or near the equatorial plane of the star? Positive answers to the last question might imply origin in disk-shaped, dissipative precursor systems. Negative answers might suggest, for example, capture of companions in dense, primordial, star-forming clusters or a high degree of asymmetry in the collapse process.

In terms of this nomenclature, one can summarize the results, as of mid-1990, of observational searches for evolved extrasolar planetary systems by stating that very-low-mass *stars* have already been found around some stars, and several discoveries of *substellar objects* have been reported—including one (HD114762b, with a mass perhaps as low as $\sim$10 $M_{Jupiter}$) that at this writing appears particularly convincing. Another body in the substellar-mass range, the $\sim$20-$M_{Jupiter}$ companion to the eclipsing millisecond pulsar 1957+20, has been reliably detected but is not included here since it was originally a stellar object that has been reduced to its present mass by proximity to the pulsar. Announced detections of *planets* have either not been confirmed by subsequent observation or are beset with such uncertainties that discovery cannot yet be claimed.

2

Basic Scientific Goals

The fundamental scientific rationale for investigation of extrasolar planetary materials has been set out in the Introduction to this report. Within this context, basic research goals may be defined for each of the broad major research areas regarded by COMPLEX as essential components of this new field of planetary exploration.

1. The search for and study of evolved extrasolar planetary systems. Scientific and technological techniques appear to be at or near the thresholds of sensitivity and precision needed to detect and study large planets in other stellar systems. The initial emphasis in this effort will focus on stars in the neighborhood of the Sun, at distances out to 10 to 100 parsecs (1 parsec = 3.26 light years = 2×10^5 AU = 3×10^{13} km) to sample an adequate population, and in the mass range of 1 $M_{\odot}$ or less because the amplitudes of stellar reflex motions are largest for low-mass stars and thus are most readily detected and precisely measured. Roughly 500 stars with a median mass of about 0.3 $M_{\odot}$ lie within 10 parsecs of the Sun. The primary initial goals are to confirm the presence and frequency of, or limit the existence of, extrasolar planets in the Jupiter- to Uranus-mass range, and to study the dynamics of discovered systems. To achieve these goals requires the following:

- *Development of instrumentation to enable sensitive search techniques, and application of these techniques to a stellar population sufficiently large to address decisively the fundamental question of existence or absence of extrasolar planetary systems;*

- *Determination of statistical distributions of occurrence, if multiple detections are made, among samples of stars of various masses and evolutionary states; and*
- *Measurement of physical and dynamical properties, including distributions of masses and orbital parameters.*

2. The study of systems of dust and gas associated with young stars and considered likely to be planet-forming environments. We know that such systems exist. Requisite observational capabilities for first-order characterization are or can be provided by existing or planned ground- and space-based observatories. Planetary materials in our own system provide some basis for interpretation of data. The general goals of this study are as follows:

- *Obtain statistics on the occurrence of dust systems among young pre-main-sequence and main-sequence stars of different types and ages;*
- *Characterize dust systems condensed as disks by determining dimensions, masses, and structural elements such as radial distributions of densities, temperatures, and orbital velocities, degrees of asymmetry, occurrence and orientation of jets, ratios of dust to gas, and broadband compositional features; and*
- *Investigate the time scales for the apparent evolution of circumstellar disks around solar-type pre-main-sequence stars from massive, optically thick structures to low-dust-mass, optically thin disks in which preplanetary or planetary bodies may have accreted or may be in the process of accreting.*

3. Supporting theoretical and laboratory studies. Real understanding of preplanetary and planetary systems requires a close interplay between new observations and theoretical and laboratory advances in areas related to the origin and physical and chemical evolution of molecular clouds, accretion disks, and planetary bodies, including the planets and accessible planetary materials (e.g., meteorites) of our own solar system. This general goal requires both new theoretical comparisons among such systems and related laboratory experiments, including considerations of the following:

- *Formation of low-mass stars, substellar objects, and planets and the differences and similarities of their formation conditions;*
- *Various types of binary systems and planetary systems and the relations among them;*
- *Fragmentation of a collapsing and rotating assemblage into multiple objects and their subsequent evolution;*
- *Physical and chemical characterization of stellar nebulae through time, silicate-carbon dust grains and icy-organic grain mantles, the precipitation of solid matter from cooling nebular gas, the conditions under which dust particles and planetesimals accrete or erode, and the evolution of accretion*

disks from initiation of dust agglomeration into subplanetary objects through subsequent planet-building stages; and

- *Definition of properties of planetary systems in the process of formation that could be subject to astronomical observation.*

3
Present Understanding of the Origin of Planetary Systems

INTRODUCTION

This section gives an overview of current theory regarding formation of planetary systems, with emphasis on relationships with astronomical observations. A great deal of effort has been devoted to the complicated problem of describing the formation of a planetary system. What is desired is the integration of a wide variety of observational evidence into a theoretical picture that includes a broad range of physical and chemical processes. Development of the theory requires calculations based on the three-dimensional hydrodynamics of a self-gravitating fluid, including the effects of pressure, viscosity, rotation, magnetic fields, shock waves, and tides. Coupled with the hydrodynamic problem are the thermodynamics of the gas and the energy transport through it, by either radiation or convection. Further, one must consider chemical processes, such as the formation of molecules and the formation, growth, and destruction of dust grains, along with their interaction with the gas. Collisions of the dust particles and their accretion into subplanetary objects, as well as gravitational and electromagnetic interactions in a many-body system, also must be incorporated. To understand the evolution of the central star in a planetary system requires the addition of nuclear physics to the above processes.

In the past 200 yr, numerous theories have been put forward regarding the origin of our solar system. Many of the earlier ideas, such as those involving capture of material from the interstellar gas by the Sun or ejection of matter from the Sun as a result of a close encounter with a passing star, have been rejected on physical grounds. The classical hypothesis of Emmanuel Kant and Pierre Simon de Laplace, that the planets originated in a disklike nebula surrounding the protosun, forms the basis for most current theoretical work on the problem. The nebula is a by-product of the stellar formation process; that is, the planets and the star are all about the same age. The committee concentrates here on the description of the problem based on the nebular hypothesis, which accounts for many, but by no means all, of the observed facts. Of course, it is possible that other planetary systems are quite different from ours in their orbital and physical characteristics; clearly the favored theory has been strongly influenced by the properties of our own system. As more information becomes available about other systems, substantial modifications of our theoretical ideas will undoubtedly be required, leading to new generalizations not yet envisioned.

Four types of observational data are crucial to understanding and potentially solving the general problem of planetary system formation: (1) general dynamical properties of our own planetary system, and statistics and properties of extrasolar planets; (2) properties of regions of current star formation; (3) statistics of multiple stellar systems; and (4) laboratory and spacecraft studies of available solar system materials (meteoritic, cometary, lunar, and terrestrial). The observed dynamical regularity of our system—the near coplanarity of the orbits, the small eccentricities and inclinations, the regular spacing of the planets, and the existence of satellite systems with similar regularities—is probably its most striking property. The masses and compositions of the inner planets as compared with those of the outer planets, as well as the existence of the asteroid belt and the composition and orbital configurations of comets, provide important clues to the nature of the solar nebula and the planetary formation process. The ordered variation in the properties of planets and satellites with radial distance from the Sun is consistent with the interpretation that they are spatially separated samples of an original continuous nebula, although accretion of each body probably occurred over a range of radial distances and the temporal sequence of formation of the planets may not coincide with their present spatial ordering. In this regard it is of great importance to measure orbital inclinations, masses, eccentricities, and other structural properties of extrasolar planetary systems.

The second type of observational information, that which describes molecular clouds and stars in the process of formation and in their early history, is extensive and diverse. It includes, for example, radio and infrared measurements. The problem is to identify those particular techniques and data that could provide clues to the origin of planetary systems. Observations of molecular clouds give some indication of the initial conditions for star formation; studies of young objects suggest the presence of disks, dense dust clouds, and mass outflows. The high spatial resolution needed to examine a nebula the size of our present planetary system even in the nearest star-forming regions (for example, at 100 parsecs, 10 AU subtends 0.1 arcsec) has been one of the greatest obstacles to progress in this area.

The third type of information relates to stellar multiplicity. A large fraction of all nearby stars of the solar type, perhaps as many as 70 to 90 percent, are members of multiple systems. There are, however, some single stars like the Sun. Whether this fraction changes significantly as a function of the mass of the primary star is observationally not well established. An important goal of the theory of star formation, therefore, is to understand the relationship between planetary formation and multiple star formation. In particular, can both occur in the same system?

Finally, data from primitive meteorites (and, ultimately, from comets) provide accurate abundance and isotope ratios for many of the chemical elements, as well as information regarding pressures and temperatures during the formation phase of these objects and therefore, presumably, of some types of preplanetary material. Measurements of both solid material and volatiles—inert gases and others—in the carbonaceous chondrites provide particularly relevant data. Some information on the magnetic fields in the primordial solar nebula can also be obtained from the meteorites, as can the time constants for some of the principal physical processes.

Four major astrophysical processes must be considered, in a unified manner, in the attempt to explain the origin of planetary systems: (1) collapse and star formation in gas and dust clouds; (2) formation, evolution, and dispersal of the disklike nebula; (3) the evolution of the central star; and (4) accretion of the nebular matter into protoplanets. The following sections discuss briefly the state of knowledge on these problems.

STAR FORMATION

The observational evidence indicates strongly that most if not all star formation takes place in molecular clouds (mean density 10^{-21} g cm^{-3}), and probably in the cores of such clouds, where densities are approximately 10^{-19} g cm^{-3} and temperatures about 10 K. The basic condition that has to be satisfied for gravitational collapse to occur is the Jeans criterion—the requirement that the thermal energy of a volume of gas be less than the absolute value of the gravitational energy. To satisfy the Jeans criterion in molecular cloud cores at this density and temperature, about 4 $M_\odot$ of interstellar gas are required. Compression to the required densities could, for example, be initiated by the passage of a shock wave through a cloud; the origin of such shocks could be supernova explosions, an expanding ionized region around an existing hot star, or the shocks associated with spiral density waves in the galaxy. However, the relatively high density cores of molecular clouds could also be formed by more gradual processes, such as the slow contraction of gaseous material relative to magnetic field lines, and a trigger may not be necessary for stars of about a few solar masses.

The simple Jeans condition neglects the effects of rotational, turbulent, and magnetic energy, all of which are significant in molecular clouds, all of which inhibit collapse, and all of which may indirectly affect the planetary formation process. Although molecular clouds rotate slowly, if the angular momentum of an element of cloud material were conserved as it collapsed into a star, the deduced stellar rotational velocity would be several orders of magnitude larger than those observed in the youngest stars. Furthermore, the increase in rotational energy as compared with gravitational energy during the collapse (again assuming conservation of angular momentum) would stop the collapse long before stellar densities were reached. The degree of ionization in molecular clouds is low but may still be sufficient for the magnetic field to inhibit collapse (magnetic braking) even if the thermal energy allows it. A further problem is that strict conservation of magnetic flux during collapse to the stellar state again predicts far-higher magnetic fields in stars than have been observed.

The density at which a protostar of about 1 $M_\odot$ begins to collapse therefore depends on the local magnetic field, but probably lies in the range of 10^{-16} to 10^{-19} g cm^{-3}. The chemical composition is similar to that of our Sun (75 percent hydrogen by mass, 23 percent helium, and 2 percent heavy elements), except that a large fraction of the heavy elements is in the form of dust grains whose size distribution and composition are assumed to be similar to those observed in the interstellar medium. The density of the material increases by a factor of 10^{15} and its internal temperature increases to about 10^6 K before it finally becomes a star. Depending on initial conditions, the collapse lasts 10^4 to 10^6 yr. Because this time is short compared with that of other stages of stellar evolution and because the protostar tends to be heavily obscured by the dusty cloud material, protostars in the collapse phase are difficult to observe. The detection of collapse velocities near the free-fall rate would provide important clues regarding the existence of protostars; instrumental capabilities are now on the verge of enabling such measurements.

Spherical collapse is an idealization of the collapse process for a system that is both nonrotating and decoupled from the magnetic field. During the early phase of spherical collapse, an isolated protostar is transparent to its own radiation, and it collapses isothermally. A very nonuniform density distribution is set up with the highest densities, and correspondingly fastest collapse rate, near the center. As densities in the central regions increase above 10^{-13} g cm^{-3}, those regions become optically thick because of the opacity of the dust grains; shortly thereafter, the collapse becomes nearly adiabatic and the interior heats up. At 1800 K molecular hydrogen begins to dissociate, and because the energy released by gravitational compression goes primarily into dissociation energy rather than into an increase in thermal pressure, further collapse is induced. Upon completion of the dissociation, the force due to the outward gas pressure gradient exceeds the force of gravity so that the collapse slows down and stops near the center. A hydrostatic core forms, and the final phase of the evolution involves the accretion of the remaining cloud material onto the protostellar core.

The foregoing discussion does not include the effects of angular momentum, which is of crucial importance in this phase of evolution. After the magnetic field decouples from the gas, the remaining angular momentum is still too large to be consistent with observed stellar rotation. Theoretical calculations indicate, however, that the collapsing cloud can fragment into two or more pieces in orbit around a common center, converting the angular momentum of the cloud into the orbital angular momentum of a binary or multiple protostellar system. The fragments can then collapse separately. Observed orbital angular momenta of wide binary systems are consistent with this suggestion. In fact, the angular momenta deduced from investigation of a number of interstellar clouds are consistent with the observation that multiple systems are often the outcome of star formation. Fragmentation may be the dominant process in the formation of multiple systems. Theoretical calculations indicate that close binary systems do not form from the fission of a rapidly rotating star in hydrostatic equilibrium. Thus although the wide binaries may form by several processes, including fragmentation, gravitational capture of one protostar by another, or the disintegration of small star clusters, the origin of close binaries is still an unsolved puzzle. Existence of a planetary system around one or more of the components of a multiple stellar system is not necessarily precluded, although the planetary orbital distances must either be very small or very large compared with the separation of the stars. Data concerning the prevalence of planetary systems associated with multiple-star systems would be of great value.

If the magnetic field transfers angular momentum efficiently from the central regions to the outer regions before protostellar collapse begins, an alternative outcome is possible. The cloud could collapse into a central star plus a surrounding disk, which would contain much of the angular momentum but relatively little mass. In this situation the disk would be stable against breakup by fragmentation, and a planetary system could form. Although the initial conditions required for the formation of a binary versus a single star are not understood, it is possible that clouds with relatively low residual angular momentum (after the magnetic braking) could take the latter course. The outcome probably depends on the distribution of density and angular momentum in the cloud just before protostar collapse begins, as well as on the total amount of angular momentum. Note that during the formation of the disk the whole inner structure is well shielded by the dust-rich outer layers of the cloud that are still in the process of collapse. Therefore at this stage the disk may not be directly observable except at radio wavelengths, but its presence could affect the infrared radiation emitted by the protosteller system. The observable layer gradually increases in surface temperature and luminosity as the central mass grows by accretion from the disk and the optical thickness of the collapsing envelope decreases. The central object becomes observable as a so-called T Tauri star, and the protostellar evolutionary phase is completed. The energetic stellar winds and bipolar outflows (jetlike collimated flows) observed during the $\sim 10^5$ to 10^6 yr of the ensuing contraction phase are now widely believed to occur simultaneously with infall of material from the protostellar cloud to the circumstellar disk and inflow of material to the central mass through the disk. There is growing evidence that these winds require the presence of an accretion disk, and that the mass inflow rate through the disk and the mass outflow rate in the wind are related.

It is important to realize that the sequence of events outlined above is simply a sketch and that only a few aspects of protostellar evolution have been calculated in detail, usually with restrictive assumptions. The general problem of collapse of a rotating, magnetic cloud, including the effects of heating, cooling, ionization, chemistry, and radiation transport, has not been solved.

PROPERTIES OF THE NEBULAR DISK

If a nebular disk forms as discussed above, it still does not have the proper angular momentum distribution to be in agreement with that deduced observationally for systems consisting of a young star plus a surrounding disk. Angular momentum must be transported out from the central regions. Various mechanisms for transport of angular momentum are under study, including gravitational or magnetic torques, viscous effects arising from turbulence or sound waves, or magnetic braking of the central star both before it appears as an optically visible pre-main-sequence star and after it reaches the main sequence. (Recent work suggests that stellar winds during intervening pre-main sequence stages are ineffective in removing stellar angular momentum.) The importance of these effects is usually studied by means of an idealized model in which the disk is thin, is small in mass in comparison with the star, is in hydrostatic equilibrium and in Keplerian rotation, and has a temperature that is low enough (<2000 K) that hydrogen is in the molecular form and dust grains are present.

Angular momentum transport by turbulence has been studied extensively, and at least two mechanisms have been suggested for inducing it. First, while matter from the cloud is still collapsing onto the disk, the difference in angular momentum between the existing disk material and the newly accreted material leads to shearing motions that could induce turbulence. Second, after the gravitational infall has stopped, convective instability can be induced by the increase in the opacity of the grains as a function of increasing temperature. The evolution of the system is therefore likely to be that of a viscous disk, often known as an accretion disk.

The turbulence has two effects: it results in an efficient transfer of heat from the midplane to the surface of the disk, and the Keplerian shear combined with the viscosity due to turbulence induces transfer of mass as well as angular momentum in the disk. The dissipation of energy by viscosity provides heat, which is radiated from the surface of the disk. The ultimate energy source for the turbulence and the heat is the gravitational energy of the disk and star interaction. Most of the disk mass eventually sinks slowly toward the central star. At the same time, the angular momentum, along with a small fraction of the mass, is slowly transferred outward. Given an infinite time for evolution, almost all of the nebula would spiral into the central star. However, time scales for circumstellar disk evolution of only a few million years are inferred from theoretical calculations, as well as from recent observations of excess infrared radiation arising from dust embedded in the disks with large infrared excesses at stellar ages of <3 million yr, but fewer than 10 percent of such stars still display this signature by the time they reach ages of 10 million yr. Such short time scales provide a strong constraint on theories of the planetary formation process, particularly for the gas-rich outer planets.

There has been considerable progress during the past few years in our understanding of the evolution of such disks. Yet major uncertainties remain in the theory of convection and turbulence that could affect the deduced evolutionary time scale. It is probable that the other mechanisms for angular momentum transport listed above could have a significant effect during the various phases of disk evolution.

The material in the nebular disk that does not condense into planetesimals or protoplanets is cleared away, arguably only a few million years or less after the formation of the star. A number of mechanisms have been proposed for accomplishing this clearing. (1) Strong stellar winds probably have sufficient energy to sweep out a moderate-mass nebula. But since massive T Tauri winds appear to be present only if the star is surrounded by a thick circumstellar disk, and are not seen in stars with low-mass, optically thin disks, it is not obvious what role, if any, is played by such winds in directly sweeping a disk away. (2) Particles in the nebula could be photoionized by ultraviolet radiation from the central star. The extra kinetic energy given to the particles, above the ionization energy, could induce pressure gradients and thereby mass loss. It is necessary, however, that nebular matter actually be exposed to such radiation, and here dust and gas opacity considerations are a major problem, particularly in the nebular midplane. One possibility is off-disk irradiation leading to evaporation of material from lower opacity disk boundaries at high altitudes. Radiation pressure on dust grains can also remove circumstellar material, but only for relatively massive stars is this process likely to be important in ejecting significant amounts of disk mass. (3) In turbulent viscosity models of the accretion disk, viscous evolution results in the spiraling of nebular material inward toward the central star, leading to stellar accretion during the protostellar evolutionary stage when inward mass flux is approximately balanced by infall of matter from the collapsing cloud and the disk is roughly in steady state. As noted above, however, there is evidence during later evolutionary phases for a relationship between mass inflow rate through the disk and mass outflow rate in massive winds. This suggests a wind-related mechanism for disk dissipation after termination of infall to the disk—ejection of most of the inwardly spiraling disk material in winds originating in the energetic boundary layer between disk and star, at rates that decrease, ultimately below current detectability, as the disk thins and inward mass flow declines. (Note that later-stage winds of this nature would not properly be called stellar since they are intercepting and ejecting mostly inflowing disk matter rather than stellar matter).

The theory of these processes has not been accurately worked out, but theoretical and observational estimates indicate that an amount of mass comparable to the primordial nebula could in principle be partly lost by outflow from the star-disk system and partly accreted within the system on time scales of a few million years, which is also the approximate time of evolution of the star itself through its classical T Tauri phase. Further detailed numerical work, in conjunction with observational and laboratory studies, is required to establish the critical time scales for angular momentum transport in the disk and for the clearing of the disk, both of which are important for the planet-forming process.

At this time, available observational evidence suggests a disk evolutionary sequence from (1) very massive 0.1 to 1 $M_\odot$ structures with high accretion rates (protostellar stage) and high mass loss rates from the star-disk system, through (2) intermediate 0.001 to 0.1 $M_\odot$ disks with intermediate accretion and mass loss rates, to (3) tenuous disks containing $<<0.001\ M_\odot$ of unassembled material with very low stellar mass loss rates. Whether a planetary system forms in all or a majority of cases within this sequence is unknown. If it does, one likely epoch of planet building appears to be during transition from stage (2) to stage (3), that is from the large infrared excesses characteristic of the classical T Tauri stars to the small or undetectable infrared excesses of the so-called naked T Tauri stars.

THE EVOLUTION OF THE CENTRAL STAR

No discussion of the origin of a planetary system is complete without consideration of the central star. The star influences the planetary formation process in several ways. First, its mass as a function of time influences the properties—for example, the vertical thickness—of the nebular disk. Second, the mass outflow from the star as it settles into the stellar state may terminate the collapse of the protostellar cloud. The observed bipolar outflows near young stars may be a manifestation of such a process. Some of the infrared objects that exhibit bipolar flow also may be interpreted as having associated disks or tori, with the plane of the disk perpendicular to the flow. As noted above, outflow from the boundary layer and ionization by ultraviolet photons may control the dissipation of the disk. Third, the tidal influence of the star plays an important role in the planetary formation process by limiting the region of gravitational influence of the protoplanets. In turn, the evolution of the nebula influences the evolution of the star, as viscous processes transfer matter to the star. The transfer of angular momentum between the star and the nebula, which also affects the evolution of the star, requires further study.

Generally speaking, once the protostellar collapse phase is over, the star is an object in hydrostatic equilibrium, with a radius a few times that which it will ultimately have on the main sequence, and with internal temperatures of a few million degrees and surface temperatures around 4000 K. The star now enters the pre-main-sequence evolutionary phase. The energy it radiates is derived from gravitational contraction, and the initial luminosity is a few times that of the present Sun. The earlier phases of the contraction proceed in the Hertzsprung-Russell (H-R) diagram along the so-called Hayashi track, that is, with nearly constant surface temperature and steadily decreasing luminosity. Energy transport in the bulk of the star is by convection during this stage and, for 1 $M_{\odot}$, the time spent on this track is about 10 million yr. The observed "classical" T Tauri stars appear within the first one-third or so of this evolutionary phase. As discussed above, they are observed to be bright in the infrared, suggesting the presence of circumstellar material. This active, classical T Tauri phase is estimated and observed to terminate within a few million years, probably through evolution by accretion or loss of disk material into the naked T Tauri stars. These T Tauri systems appear to be very likely sites for accretion of subplanetary and protoplanetary objects.

After the Hayashi phase, the path followed by a contracting star changes direction in the H-R diagram, evolving with gradually increasing luminosity and increasing surface temperature until hydrogen burning ignites at the center. During this phase the energy transport within the star is primarily by radiation. When nuclear reactions contribute 100 percent of the energy output, the star is said to have arrived on the "zero age" main sequence.

Some aspects of the evolution of the star seem to be reasonably well understood. Others, such as the role of its rotational and magnetic energies in generating and collimating bipolar outflows, still require further and more detailed observational and theoretical study.

FORMATION OF THE PLANETS

The theory of planetary formation rests on calculations of particle dynamics, collisional accretion theory, and gas dynamics, some aspects of which are still quite uncertain. The committee discusses briefly the concept that all planets formed by essentially the same process, that is, by gradual accretion of small dust particles into larger subplanetary bodies (commonly called planetesimals in this context), which later coalesced to form the planets. This scenario, although not the only possible planetary formation process, is one that is now undergoing intensive study. Even though models based on this picture are oversimplified because of existing limitations of computers, they contribute to the advance of general insight. In these models, the major difference between the inner and outer planets is simply that the latter grew to the point (10 to 20 $M_\oplus$)where they were able to attract a significant amount of nebular gas to form an envelope around the solid core, which presumably consisted of both rocky and icy material. In the case of Uranus and Neptune the gaseous envelope was much smaller in mass compared to the rest of the planet, and part of the accretion of solid matter could have occurred after dissipation of the nebular gas. In the inner solar system both the heating and tidal effects of the Sun, as well as the smaller amount of condensable material, apparently prevented the buildup of cores to the critical mass where significant gas accretion was possible.

The starting point is the nebular disk, composed of gas mixed with about 1 percent by mass of dust, at temperatures in the range of 100 to 2000 K. The dust particles have essentially interstellar characteristics, with typical particle sizes of 10^{-5} to 10^{-4} cm (0.1 to 1 μm). Dust would be absent close to the central star, where temperatures are expected to be high enough to vaporize it. In the inner parts of the nebula beyond this region, out to the point where the temperature falls to roughly 200 K, the particles are composed principally of compounds of oxygen, magnesium, silicon, and iron. In the outer regions, below 200 K, water ice can also exist, as well as ices of ammonia, methane, and various clathrates.

The first stages of dust accumulation into larger objects are proving to be perhaps the most difficult of all the phases of planetary formation to understand. In one long-standing scenario, dust grains gradually sink to the midplane of the nebula, growing by accretion to centimeter size on a time scale of a few thousand years. Once the dust layer at the midplane becomes dense enough, gravitational instability occurs. The layer fragments into rings, and these further fragment into gravitationally bound aggregates with characteristic sizes of a few kilometers and masses on the order of 10^{18} g at the Earth's distance from the Sun. This gravitational instability model of planetesimal formation is simple and appealing, but there are now serious questions concerning many of its basic assumptions and predictions. They focus in particular on the disruptive effects of turbulence in nebular gas, the role of stickiness of dust grains in grain coagulation processes, the physical morphology and settling times of resulting dust aggregates, and the short time scale of $\lesssim 10^4$ yr predicted for accretion of kilometer-size objects (there are estimates that it may have taken 10 to 100 times longer). It is fair to say that at present the mechanisms of growth to this size range, and their required time scales, are poorly understood.

There is, however, a widely accepted standard model for subsequent planetary accumulation. Planetesimals, whatever their means of formation, undergo collisions and gradually accumulate into a few large bodies that eventually form the terrestrial planets and the cores of the giant planets. Monte Carlo simulations of the accumulation process suggest that it proceeds rapidly during the first few million years, forming objects up to $\sim$25 percent of the Earth's present mass in the terrestrial planet zone; subsequent collisional growth occurs more slowly because accreting objects decline in number and become orbitally more isolated from the protoplanets. Estimates for the total time required to fully accrete the terrestrial planets range from 10^7 to 10^8 yr. Once a system of planetesimals forms, it also serves as a reservoir from which dust can be eroded over much longer time scales. As discussed further in Chapter 4, collisional processes could explain the observed maintenance of dust around main-sequence stars 10^9 yr or more after formation.

The properties of our own solar system strongly indicate that Jupiter must have formed more rapidly, because (1) the core must have accreted to its present size before the nebula gases disappeared and (2) the presence of the asteroid belt without a major planet in it strongly suggests that the prior presence of Jupiter and its gravitational influence prevented the final stages of accretion from occurring there. Current research is directed toward the question of how to build Jupiter's core, which probably contains about 20 $M_{\oplus}$, within a few million years (if, in fact, this correctly represents the time scale for dissipation of nebular "gas." There are no observational astronomical constraints on the lengths of time required for disappearance of the gaseous component of accretion disks; they must be inferred from the dust survival times deduced from infrared excesses, and this may not necessarily be a valid extrapolation.). It is possible that improvements in the theory will show that the phase of rapid accretion of planetesimals continues to well beyond lunar mass at Jupiter's distance from the Sun. If that is true, Jupiter's core could build up quickly to the point where it could start accreting gas.

An alternative theory, in which all planets formed as large gaseous condensations, of which only the cores persist in the case of the terrestrial planets, has encountered a number of difficulties. In particular, the high interior pressure may suppress bonding differences, making the precipitation of core material unlikely. This idea does provide a natural way of forming the giant planets quickly, but a considerable amount of theoretical work is necessary to clarify the formation process of these objects.

An understanding of the early evolution of the giant planets is of particular importance because of the possibility that they can be detected near young stars. A number of observational constraints can be used to guide the theory, including the present chemical composition of the giant planets, their present masses, luminosities, radii, and gravitational moments, and the fact that all except Neptune have regular satellite systems.

Detailed numerical calculations support the following scenario for the various phases of the evolution of Jupiter. First, the solid core builds by accretion to about 1 $M_\oplus$. Second, the gravitational influence of the core attracts nebular gas, which forms a thin layer in hydrostatic equilibrium around the core. As the core continues to accrete planetesimals, its region of influence (defined by its accretion radius or tidal radius, whichever is smaller) grows, causing further gas accretion. Third, the core, having approached the critical mass, has sufficient gravitational influence to attract an envelope of comparable mass. The energy radiated by the envelope then can no longer be supplied by the accretion energy of the planetesimals, and the envelope begins to contract. Fourth, the envelope mass increases rapidly, reaching its present value in a few thousand years, with no significant change in the core mass. The radiated luminosity increases also, because of the rapid contraction and increasing mass, to values of 10^{-4} or 10^{-3} times the present solar value, far higher than the luminosities (10^{-9} to 10^{-11} $L_\odot$) of the present giant planets. Fifth, gas accretion terminates. The termination may result because the protoplanet has grown to a mass sufficiently large that its tidal influence on the nebula opens up gaps in its vicinity and prevents further accretion of gas onto it. Alternatively, the accretion could terminate when there is no gas left to accrete after escape of gas from the nebula.

Thereafter, the protoplanet evolves at constant mass and contracts on a time scale of 10^5 yr, releasing energy at a level of 10^{-4} $L_\odot$ at surface temperatures of a few thousand Kelvin. This phase could possibly be observable. The later evolution involves cooling of the interior and only a very slow decrease in radius; the luminosity and surface temperature decline as a function of time. After 4.5 times 10^9 yr the present radius and intrinsic luminosity of Jupiter are reached.

The above theory of the evolution of the giant planets is based on numerous approximations, does not include rotation and the formation of satellite systems, and will undoubtedly require extensive revision in the future; nevertheless, it agrees with many known properties of the giant planets. For example, the theoretically deduced critical core masses are in good agreement with the core masses deduced from observations of the giant planets. The formation process for the outer planets Uranus and Neptune, however, is still not well understood. In particular, the time scale currently inferred for accretion of their cores at their present distances is longer than the lifetime of the nebula and may be longer than the age of the solar system. Perhaps resonances played some role, as suggested by such observations as the trapping of Pluto in a 3:2 resonance with Neptune, and the extraordinary circularity of Neptune's orbit. Moreover, it is not understood how these planets could have acquired their gaseous envelopes of approximately 1 $M_{\oplus}$. Possibly the accretion of Uranus and Neptune may have started closer to the Sun than their present distances. Their cores may have grown sufficiently large to attract some gas from the nebula, and later the protoplanets could have migrated to the outer parts of the nebula under the gravitational influence of Jupiter and Saturn. Their buildup could have been completed simply by the accretion of planetesimals.

What is certain is that various critical aspects of planetary formation still need to be clarified. Progress depends on development and testing of improved models for the process, and thus on continuing theoretical and observational investigation of planetary systems. Many theoretical and computational initiatives will require expansion of currently available computer capacity and improved numerical techniques for solving coupled systems of differential equations. New and detailed observational data on chemical compositions, present physical states, and dynamical behavior both within and outside our own solar system, from spacecraft instruments and from ground and Earth-orbital facilities, are central to this effort.

4
Observations of Extrasolar Planetary and Protoplanetary Material

INTRODUCTION

No widely accepted, confirmed discoveries of extrasolar planets ($M \leq 2\ M_{Jupiter}$) have been made as of the writing of this report, but searches are continuing with techniques that currently are capable of detecting high-mass planets in favorable orbits. There is, however, abundant evidence for a variety of planetary and preplanetary materials associated with stars, pre-main-sequence stars, and what may well be protostars. Present observing programs are likely to augment this evidence rapidly, as it is generally true in astronomy that the discovery of a new type of object or phenomenon presages the identification of a new class.

This chapter gives the status of current research on the observable properties of extrasolar planetary materials. These include possible planets and substellar objects, of course, but also gas and dust around stars, both in their formative phases and in their main-sequence lifetimes. Dust created and ejected into interstellar space by stars during their later stellar evolution is of interest as well, since it constitutes one type of raw material for subsequent generations of star formation. The committee examines each of these topics in the following sections.

MATERIAL ASSOCIATED WITH PROTOSTARS AND PRE-MAIN-SEQUENCE STARS

The observable properties of protostars and pre-main-sequence stars, collectively known as young stellar objects, are very different from those that characterize stars in their later life. Disklike structures and collimated bipolar outflows of material are seen in the pre-main-sequence phase, evolving on rapid time scales to the much more tenuous concentrations of circumstellar matter observed in somewhat older pre-main-sequence systems. The distinctive appearance of young stellar objects is associated with their evolutionary properties, which involve the contraction of a star toward a stable main-sequence configuration (see The Evolution of the Central Star in Chapter 3). The order of events and their manifestations are dependent on mass. Pre-main-sequence stars of low to moderate mass, defined here as less than 3 $M_{\odot}$, go through the T Tauri phase. T Tauri stars are very young; the optically thick circumstellar structures that characterize this evolutionary stage are observed at inferred stellar ages ranging from <3 million yr up to roughly 10 million yr. Since the range of survival times for these structures is only a small fraction of their central star lifetimes, such objects are relatively rare.

T Tauri stars are characterized by a variety of phenomena, many of which distinguish them from ordinary main-sequence stars: (1) emission from hot (1000 to 1500 K) dust (termed "an infrared excess" because the strength of the emission exceeds that of the underlying stellar continuum); (2) large optical, infrared, and x-ray variabilities on a variety of time scales; (3) 10-μm silicate emission features, occasionally observed in absorption; and (4) occasional strong linear polarization in the optical and near-infrared continuum. The variable infrared excesses do not correlate with the amount of observed visible reddening (due presumably to circumstellar dust absorption), implying that the variation is intrinsic to the star. The variability is probably ultimately due to active regions (starspots) on the young stellar objects. These active regions are presumably regions of intense magnetic field activity—generated by vigorous convection and, possibly, by differential rotation induced by magnetic braking. Hot dust is sufficient to explain the infrared radiation from T Tauri stars; whatever hot gas is present is not dominant.

According to current thinking, very young pre-main-sequence stars with a wide range of stellar masses are thought to be associated with a class of objects characterized by strong collimated outflows of gas and energetic interactions with the molecular clouds in which they are embedded. These regions of interactions, known as Herbig-Haro objects, are seen both optically and in carbon monoxide line emission. The presence of disks has been inferred for stars covering the mass range from 0.2 to ~5 $M_{\odot}$.

Evidence of molecular outflows are found for both optically obscured sources, still located within their protostellar cores, and visible sources throughout this mass range. Stars more massive than $\sim$5 $M_\odot$ have not been observed optically in the pre-main-sequence phase, presumably because of their short contraction times. Instead, at this stage, they are seen only as luminous infrared sources, embedded near the centers of molecular clouds.

The characteristics and properties of disks associated with young stellar objects are studied by a variety of observational techniques, each sensitive to particular aspects of disk structure. In general, disks actually appear as elongated or barlike structures that are interpreted to be disks in projection. In only a few cases do molecular line measurements indicate rotation appropriate to a disk or torus. In other cases, the presence of a disk is largely based on models of bipolar outflow, with disks and bipolar outflow formation being causally linked, or on observations of one outflow direction being blocked by a disk or a disklike object.

At the largest scales, disks are apparently manifested as elongated molecular clouds $\sim10^3$ to 10^4 AU across. Molecular hydrogen number densities (n_{H_2}) are estimated to be $\sim10^3$ to 10^4 cm^{-3} at a typical radius of $\sim10^4$ AU. The degree of inferred flattening of these disklike structures is not large; the ratio of the hydrogen column density along a radius parallel to the elongation to that perpendicular to it (the axial density ratio) is typically only a few. The inferred flattening is considered reasonably good evidence for rotational or magnetic support of a disk, although flattening by itself is not robust evidence for either. While disks of this scale are clearly larger than those usually thought of as circumstellar disks (radii $\leq$1000 AU) or inner accretion disks from which planets may be born (radii $\leq$100 AU), it is reasonable to assume that these smaller-scale structures are embedded in the larger-scale ones, especially as the initial collapse phase may occur in a larger-scale cloud. In fact, microwave continuum observations show that within the larger disk structures, smaller (radii $\sim$2,000 AU) elongated structures are embedded with the same orientation. Molecular hydrogen abundances are increased, with n_{H_2} densities of at least 10^6 cm^{-3} in the plane. These structures are presumably more flattened, with axial ratios of $\sim$10 to 100.

Inner, circumstellar "disks" or elongations can be indirectly detected by observing infrared radiation scattered from young stellar objects. Speckle interferometry and maximum entropy techniques have been used to infer the existence of disks in the approximately 100- to 1000-AU radius range. We anticipate that use of infrared array cameras will be of value in detecting scattering disks toward young stellar objects.

Disks can be directly detected at radio wavelengths. Radio continuum and millimeter-wave observations are sensitive to thermal emission by dust in disks, and molecular line observations are sensitive to emission and absorption by gas. HL Tau is a particularly intriguing object for study by these methods. Here, the total amount of gas and dust lying in a rotating disk is inferred to be as high as 0.1 to 1.0 $M_{\odot}$. Thus it is possible that a small disk with dimensions comparable to those of our planetary system may contain sufficient matter to build a system similar to our own.

Molecular outflows associated with young stellar objects, detected by means of radio emission lines, characteristically extend for $\sim 10^4$ to 10^5 AU. These aligned, conical structures often have collimated infrared and optical manifestations as well. Carbon monoxide flow velocities are typically 10 km s^{-1} ranging up to 100 km s^{-1}, but velocities in highly collimated optical jets (often linear chains of Herbig-Haro objects) range up to several hundred kilometers per second. Estimated lifetimes of the molecular outflows are between 10^4 and 10^5 years, a small fraction of the duration of the T Tauri phase. The mechanical luminosity of these flows is significant compared to the bolometric luminosity of the young stellar object, implying a driving mechanism far more powerful than radiation pressure or ordinary stellar coronal winds. Some of the most significant observations are geometric: collimated outflows of neighboring stars are often aligned with each other, aligned with cloud or disk magnetic fields, or oriented perpendicularly to the long axis of circumstellar disks or molecular clouds. The opening angles of the outflows are generally less than 45° and become smaller with increasing resolution of the source area. Also, blue-shifted, collisionally excited emission lines of the optical outflows are seen in preference to red-shifted ones; the inferred red-shifted companion flow is believed to be obscured by dust. The central pre-main-sequence star can be obscured as well, possibly implying the presence of a disk or torus.

The observations above fit within the theoretical framework of a young stellar object surrounded by a circumstellar disk. The existence of a central accretion disk of ~100-AU scale is inferred, but no certain example has been directly resolved and studied in detail. Cloud magnetic fields may determine the orientation of the rotation axes of stars and disks. All current theoretical models of the mechanism of bipolar outflows invoke an accretion disk, either in an active role of providing kinetic energy for the flow or in a passive role as a barrier to a more isotropic flow whose energy is provided by the young stellar object itself. When these disks are seen edge on, the radiation is linearly polarized; some researchers have inferred a magnetic field oriented perpendicularly to the disk. HL Tau and DG Tau are edge-on examples.

In general, newly forming substellar objects or planets within these disks cannot be seen due to inadequate observational sensitivity. One exception is T Tauri itself. It has an infrared companion, ~80 AU distant. Based on very large array (VLA) radio measurements, both objects have powerful ionized stellar winds, but the companion dominates the T Tauri outflow. One estimated mass of the T Tauri companion is substellar, and its luminosity can be interpreted to result from capture of matter from an accretion disk. Although further infrared and radio searches for substellar or protosubstellar companions to T Tauri stars have not been definitive, such searches are actively under way and have led to some tantalizing results. The existence of substellar or even planetary companions in the process of formation does not appear to be an unreasonable supposition.

Observations of disks and jets around young stars stand as interesting and largely unexplained phenomena. And as detailed in Chapter 7, they may have some connection to the physics involved in collimated bipolar outflows of drastically different scales and energies, associated with such objects as SS433, Seyfert galaxies, active galactic nuclei, and quasars.

GAS AND DUST DISKS AROUND MAIN-SEQUENCE STARS

As young stellar objects evolve onto the main sequence and begin to age, it is anticipated that they shed their associated disks of matter and become essentially unobscured. The naked T Tauri stars observed in association with the generally younger classical T Tauri stars seem very likely to represent just such an evolutionary stage of disk dissipation in these stars. It is therefore of considerable interest that dust and gas have been detected around at least some moderately evolved main-sequence stars. A central question that these observations raise is whether we are observing the remnants of a primordial circumstellar dust and gas disk, and thus the remnants of an accretion disk. One plausible class of remnants is planets.

Many of these circumstellar dust observations were made by the IRAS, which detected large infrared excesses associated with what had appeared to be ordinary main-sequence stars. In particular, higher-resolution, ground-based infrared measurements had not detected dust close (<100 AU) to these stars. First reported was an extended region around Vega. Since then, dozens of stars have been classified as Vega-like on the basis of their 60-μm and 100-μm flux excesses, although spatial information is not generally obtainable. About half of these stars are spectral type A (1.5 to 2.0 $M_{\odot}$), but excesses are also seen in more solarlike F, G, and K stars (0.5 to 1.5 $M_{\odot}$). Given the likely selection effects caused by the volume and flux limitations of the survey, it is entirely possible that *most* main-sequence stars have such dust clouds. We may have only detected the thicker, more massive ones. Further discoveries will better define how cloud properties depend on stellar properties such as spectral type, duplicity, age, rotation, metal content, and so on.

Three stars for which spatial information is currently available are Vega (αLyrae), Beta Pictoris, and Fomalhaut (αPsA).

1. *Vega*: the dust source extends to a radius of ~80 AU, perhaps as far as 250 AU. The size of the dust particles is ~80 μm based on infrared emission characteristics, and the dust has a color temperature of 85 K. Thus it is possible that these dust grains may be icy. The ~80-μm particle diameter is some 2 orders of magnitude larger than typical interstellar grain sizes of 0.1 to 1 μm, indicating that the dust is not primary (i.e., interstellar) and that local coagulation-condensation processes have occurred. The minimum mass of the dust region is estimated to be $\sim 10^{25}$ g ($\sim 10^{-3}$ $M_\oplus$), but much more could exist in larger bodies. No gas absorptions are seen in the visual or ultraviolet.

2. *Beta Pictoris*: optical charge-coupled device (CCD) images show an elongated source extending to beyond a 500-AU radius in two opposing directions. Modeling based on visual coronographic and infrared studies suggests a disk inner edge at ~17 AU. The minimum dust mass to account for the observations is $\sim 4 \times 10^{25}$ g ($\sim 10^{-2}$ $M_\oplus$); again, much more mass could exist in larger bodies. Remarkably, narrow calcium and sodium lines are seen in visual and ultraviolet absorption; there is apparently a gas disk that accompanies the infrared and visible dust disk. The mass of this gas is estimated at <2 $M_\oplus$, with a characteristic $n_{H_2} \sim 10^5$ cm^{-3}.

3. *Fomalhaut*: material is detected out to a radius of ~140 AU and possibly beyond. Certain interstellar gas absorptions may actually be circumstellar.

It is plausible that all three of these systems have gas disks, but Vega and Fomalhaut lack absorptions because they are not seen close to edge on. Details of other detections continue to be reported in the literature.

A number of questions need to be addressed before astrophysicists can conclude that any of these stars has planets:

- Do we observe gas and dust because these systems are young (the lifetimes of A-type stars are $<10^9$ yr) and are thus direct descendants of a preplanetary (protosolar) nebula? Recent data suggest that a number of old ($>10^9$ yr) main-sequence stars may have disks.
- Does the absence of obscuration and emission from hot dust within the central regions of some of the systems imply that these regions have been cleared of dust? Relatively sharp edges on spatially resolved disks imply some clearing process.
- Could this clearing process be due to planetesimal accretion or to a more prosaic phenomenon such as gas drag?
- Do the observed dust disks require replenishment from an invisible swarm of larger bodies?
- And could type A stars go through an evolutionary phase, unrelated to the standard protostellar and pre-main-sequence phases, that results in the formation of equatorial shells of gas and dust? Rapid removal rates

for dust strongly suggest that renewal is required. A recent interpretation that such disks are resupplied by relatively large, but subplanetary, bodies (perhaps analogous to asteroids or comets) and truncated at the inner edge by planets, has been advanced based on analyses of IRAS data.

SUBSTELLAR OBJECTS

The preceding section emphasized the detection of dust and gas around early-type (and thus luminous) main-sequence stars. Such observations are currently much more difficult around the more common, faint late-type dwarf stars unless they are close, and will probably remain so until the Space Infrared Telescope Facility (SIRTF) deployment brings objects of this kind, even at distances of hundreds of parsecs, into effective observational range. Nevertheless the possibilities for detection of massive planetary and substellar companions around faint low-mass stars are enhanced, both because of more favorable secondary to primary luminosity ratios, and because astrometric wobbles and radial velocity Doppler shifts due to stellar reflex motions are greater for a given companion mass and separation. These topics are treated in greater detail in Chapter 5.

Ongoing astrometric and radial velocity work points to the existence of companions ranging in mass from small stellar to substellar. Near-infrared speckle interferometry has generally confirmed the presence of low-mass stellar companions with separations of >0.06 arcsec. Of the substellar companions reported, one (HD114762b, $M \gtrsim 10\ M_{Jupiter}$) appears at this writing to be relatively solid. Claims of planetary companions are considerably more uncertain at the present time. Several of the unconfirmed companions have predicted separations of <0.06 arcsec, and some would not necessarily be luminous enough in the infrared to be directly detected in any event. It is also quite possible that many of the weaker astrometric perturbations are not real.

The first claimed direct detection of a substellar mass object, a companion (VB 8B) to Van Biesbroeck 8 (VB 8), was made in 1985 by two-color infrared speckle interferometry on three occasions by one observational team. Theoretical bounds on its mass ranged between ~40 and 80 $M_{Jupiter}$. More recent observations have, however, failed to confirm the existence of VB 8B. Near-infrared speckle interferometry has also been applied to nearby stars not suspected of having astrometric companions. None were found, and the limits on companion luminosity generally preclude objects resembling the putative Van Biesbroeck substellar body.

The controversy surrounding VB 8B illustrates the great difficulty of directly observing such faint objects with current techniques. Because substellar objects more luminous than VB 8B occupy a rather narrow regime in age-mass space, definitive identification of such objects, either as binary companions, components of multiple-star systems, or perhaps as isolated objects, will continue to be a tremendous challenge over the next decade. With the exception of the very recent report of substellar bodies in Taurus, at the time of this writing, only three other candidates—Gliese 569B, GD 165B, and LHS 2924—have emerged from imaging or spectroscopic observations. The first two are cool companions to white or red dwarf stars, all have possible masses as low as ~50 $M_{Jupiter}$ or so, and none are considered proven as yet. LHS 2924 is cooler and less luminous than any known star, but it has a peculiar spectrum, which hampers its positive identification. Deep CCD integrations in the red and near-infrared in limited areas have turned up no examples.

Data from continuing searches for indirect stellar reflex effects (astrometric wobble and Doppler shifts in the velocity of the orbital reflex, as opposed to direct imaging or spectroscopy) are encouraging in a few cases, and in general leave open the possibility that substellar objects (and perhaps planets) may exist around other nearby stars. More measurements over a longer time period are needed to clearly confirm—or refute—claims of detection.

CIRCUMSTELLAR DUST

It is appropriate at this point to consider the properties of circumstellar and interstellar dust in greater detail. Circumstellar dust has been observed as absorbing, reflecting, or emitting matter that surrounds cool giant and supergiant stars, planetary nebulae, hot stars, evolved stars, novae, supernovae, and pre-main-sequence stars. Most of the dust around pre-main-sequence stars is probably preexisting interstellar material, but for the majority of other stars mentioned above, where they are observed, circumstellar dust shells appear to be condensates formed in situ from gas derived from the central star. Circumstellar dust is important to studies of planetary systems because it is the original source of interstellar grains, and because close examination of dust shells may give detailed information on condensation processes and products that occur in gases of various temperature, density, and composition. At the present time, however, it is not clear what direct links there might be between grains observed to form in circumstellar envelopes and particles that eventually accrete to form planetary materials. Evolutionary processes that occur in the interstellar medium may drastically alter interstellar grains between the time they leave, for example, the atmosphere of an M giant and the time they are incorporated into a solar nebulalike environment where planetary bodies could form. Current models of grain sputtering by shock fronts indicate that grain destruction may occur on a time scale that is a factor of 10 shorter than the billion-year period required to replenish the interstellar medium with dust observed to escape circumstellar envelopes. If this is actually the case, then dust must reform in the interstellar medium.

The existence of circumstellar dust is most often revealed in the infrared. The dust is heated by radiation from the central star and, in cases where active disk structures are present, possibly by accretion of material through the disk. The observational evidence consists of excess infrared radiation from the heated dust above the underlying stellar continuum, or of spectral features due to silicates, ices, or carbon-rich materials. In the visual range, dust leads to reddening, extinction, and polarization or can be directly imaged in reflection nebulae. The infrared excesses range from a small change of continuum slope, usually beyond 5 μm, to cases where the infrared flux far exceeds the observed visual luminosity. The magnitude and spectral energy distribution of the excess can be used to estimate the dust mass surrounding the central star, and to diagnose dust temperature distributions and optical depth structures.

For late-type giants the total mass of the dust envelope correlates with the gas-loss rate determined by radio carbon oxygen measurements. The ratio of gas to dust implied from these measurements is $\sim$200 to 300, indicating that virtually all condensable refractory elements have formed grains. At least for late-type giants, condensation from outflowing gas seems to be a very efficient process. This is confirmed by the remarkable efficiency of condensation actually observed to occur during nova outbursts.

Information on the composition of grains can in principle be determined from the temperatures of their formation and stability. In practice this is complicated by kinetic effects, nonlocal thermodynamic equilibrium environments, and the complexity of gas outflow. Infrared spectral features provide rather direct information on grain composition. The 9.7-μm and 18-μm features seen in emission and absorption around oxygen-rich stars have been attributed to the stretching and bending vibrational modes of silicon-oxygen bonds in silicates. The shape of the "10-μm feature" is not consistent with that of single-phase crystalline silicates that have been studied in the laboratory, and the suggestion has been made that the circumstellar silicates are amorphous or poorly ordered. In rare systems, a 3.1-μm feature is seen that is consistent with water or ammonia ice at temperatures of less than 150 K.

Around carbon-rich stars where the ratio of carbon to oxygen is greater than 1, the grain chemistry is quite different. In these environments, nearly all of the gas-phase oxygen is bound up in carbon oxygen, and the excess of carbon, coupled with the lack of water vapor, favors the formation of carbon-rich and reduced grains while depressing the formation of silicates. In the carbon-rich stars a broad feature at 10.5 to 12 μm has been identified as stretching-mode vibrational emission by silicon carbon and another feature near 25 μm has been associated with magnesium sulfur. Because the silicon abundance is small compared to that of carbon, silicon carbon cannot be the dominant grain in these systems. The major part of the infrared excess in carbon stars is probably due to fairly pure carbon in either amorphous or crystalline form. The apparent absence in the observational spectra of a predicted resonance feature at 11.5 μm indicates that the carbon is not graphitic. The previously unidentified infrared emission features at 3.3, 6.2, 7.7, 8.7, and 11.3 μm in carbon-rich outflows have recently been attributed to emission by partially hydrogenated polycyclic aromatic hydrocarbons (PAHs). The PAHs may in addition be responsible for other infrared features and may be the source of the diffuse interstellar bands seen in the visible.

Future efforts from a variety of disciplines are needed to determine possible links between circumstellar dust grains and materials found in meteorites and interplanetary dust. Although no individual grains that have been studied in the laboratory have been shown to be interstellar or circumstellar in origin, there is fairly strong isotopic evidence that materials from circumstellar envelopes did survive transport to the solar nebula and were incorporated into some meteorite parent bodies. The most direct evidence for this is the survival of nearly pure ^{22}Ne (neon-E) in carbon-rich separations from carbonaceous chondrites. If this exotic neon component is the product of the decay of ^{22}Na (half-life = 2.6 yr), then the sodium must have been incorporated into solid grains within a few years of its synthesis. New infrared observations of novae suggest that both ^{22}Na production and grain formation may be observed in binary novae involving matter accreting on an oxygen-magnesium-neon white dwarf. The neon-E component is the most obvious example of probable survival of circumstellar dust and incorporation into the solar system, but many of the other isotopic effects in meteorites may have had a similar origin. Heavy carbon, light nitrogen, and s-process krypton and xenon are also found in carbon-rich separates. These isotopic effects are consistent with formation from a material derived from a red giant. The apparently correlated effects in the neutron-rich equilibrium isotopes ^{48}Ca, ^{50}Ti and ^{54}Cr may also be tracers for circumstellar material ejecta from supernovae. Large deuterium/hydrogen enrichments, a factor of 3 or larger, have been seen in carbonaceous chondrites, in the organic fraction of unequilibrated ordinary chondrites, and in interplanetary dust. No known solar system process can produce such large isotopic fractionation in hydrogen, but similar and even larger effects are seen in molecular clouds, where they are believed to be produced by ion and molecule reactions. The heavy hydrogen in meteoritic materials is believed to be a direct link to the chemical evolutionary process that occurred in molecular clouds. This type of information has important implications for the evolution of biogenic elements and their incorporation into planetary materials.

5
Observational Requirements for Identification of Extrasolar-system Planets

OBSERVATIONAL HISTORY

The observational search for extrasolar-system planets is rooted in studies of binary and multiple-star systems in the solar neighborhood. Knowledge of the stellar mass distribution in these systems, especially in the large preponderance (>50 percent) of binary stellar systems, is important to the theory of star formation. Many of the companion stars are unseen and can be detected only indirectly by measurement of stellar reflex motion.

Traditionally, astrometric measurements of the motion of nearby stars have been based on data recorded in plate collections spanning several decades. Studies of this nature have produced reports of purported planetary-mass companions to nearby stars, the most famous example being Barnard's star. One interpretation of the astrometric measurements for this nearby stellar system (the second closest to the Sun) proposes the existence of two sub-Jupiter-mass planets with periods of 12 and 20 yr; this interpretation, however, has been questioned on the basis of independent observations. Based on astrometric data a substellar-mass companion, VB 8B, was also proposed for Van Biesbroeck 8 and, as noted in Chapter 4, a putative substellar object was reported to have been detected directly by speckle interferometry but was not found in subsequent infrared imaging. Moreover the astrometric data that initially suggested the presence of a substellar-mass companion to VB 8 are now being questioned by later photoelectric astrometric measurements.

Recent studies on precision radial velocity measurements suggest the possible existence of orbiting bodies around a small number of nearby stars, one of which (γ Cephei b) is inferred to have a mass perhaps as small as that of Jupiter. These detections have yet to be confirmed. Although, as noted earlier, the data for HD114762b are promising, and Gliese 569B, GD 165B, and the very recently announced Taurus objects appear to be viable candidates, at present there is no conclusive evidence from any observational technique for even a single star with a substellar-mass companion, and there is considerably more uncertainty regarding current claims of discovery of planetary-mass bodies. It is with this background in mind that the committee examines the observational requirements for an extrasolar-system planet search.

The scientific value of searches for new planets lies in knowing the statistics of their occurrence and their locations as a basis for continuing physical studies. These benefits decline with uncertainty about mass; that is, with ambiguity as to whether the orbiting object is, in fact, a planet. The benefits increase in proportion to the proximity of the discovered system, which translates into greater accessibility for further research. Beyond proving the "existence theorem" by discovery of planets, each potential search technique provides some physical information about the discovered objects—more or less enmeshed with ancillary assumptions required for interpretation. In the next section, the committee discusses possible approaches to direct detection of substellar- and planetary-mass companions.

DIRECT DETECTION

Any direct-detection system consists of a telescope coupled to ancillary instrumentation that analyzes and records the collected radiation. The efficiency with which a telescope performs its functions of gathering, relaying, and focusing radiation is crucial because of the very low intrinsic brightness of extrasolar planetary material. The committee considers for illustration the challenge of observing Jupiter from a nearby star, using two space-borne telescopes in NASA's astrophysics program: the Hubble Space Telescope (HST) and the proposed Space Infrared Telescope Facility (SIRTF). The committee also considers the limitations on direct detection of a Jupiter-size planet in orbit about another star. Jupiter is selected for this analysis because its brightness and relatively large distance from the Sun optimize the chances of direct imaging.

The HST and SIRTF are satellite observatories designed to have diffraction-limited optical quality. HST has a 2.4-m aperture and a complement of cameras and spectrographs for recording the ultraviolet and visible light typically generated by stars. SIRTF is expected to have a 0.85-m aperture and to operate in the infrared from 2 to several hundred microns. Dust, rocks, or planets gravitationally bound to a star and warmed by its radiation would emit thermal radiation in SIRTF's spectral range. Indeed, the dominant noise background for most SIRTF observations would come from solar system zodiacal dust through which the telescope must view astronomical targets in deeper space. In principle, a "Jupiter" might be directly detected by either HST or SIRTF, in reflected starlight or in emitted thermal radiation, respectively.

The dominant noise for such detections arises from the portion of the stellar image that overlies the nonstellar signal. The stellar background at any angular distance from the center of the telescopic image of the star is determined by the diffraction pattern of the telescope pupil and by scattering from the telescope optics. There is a strong wavelength dependence for both the diffraction image and the energy distributions of the star and the planet.

These points are illustrated by the following analysis. There are no stars closer to the Sun than 1 parsec, but there are several hundred within about 10 parsecs. For realism the committee considers the problem of detecting Jupiter in our solar system from an intermediate distance of 5 parsecs. At its greatest elongation from the star, the apparent angular distance to the planet would be 1 arcsec, and its flux at 0.25 μm wavelength about 10^{-9} times that from the star. At 20 μm, where Jupiter's thermal emission peaks, the flux ratio is 10^{-5}. Jupiter's 10^{-5} to 10^{-9} contribution to the bulk light would be undetectable as an incremental light intensity and thus rules out detection without spatial discrimination.

Consider the limitation on spatial resolution imposed by the optical properties of the telescope. For diffraction-limited optics, the dominant component of the HST and SIRTF point-spread functions—the telescopic image of an unresolved astronomical source—is an Airy diffraction pattern with the first dark ring at radius 0.026 arcsec for HST at 0.25 μm wavelength, and at 5.9 arcsec for SIRTF at 20 μm. In the latter case there is no hope of distinguishing Jupiter's image because its separation from the Sun is much smaller than the image size. In the future, direct detection of planets in the infrared, making use of the more favorable contrast than at visible wavelengths, may be possible using interferometric techniques. Moreover one should also note that Jupiter radiates more energy than it absorbs from the Sun (as do Saturn and Neptune as well), and a Jupiter-like object could be observably self-luminous in the infrared if it were located >5 AU from a Sun-like primary. Assessments of detectability become considerably more complex than that given here when this property is included, but potential observability at large orbital distances clearly increases, and searches for companions in the mass range of the giant solar system planets are planned to be a major scientific component of the SIRTF observing program.

For HST the image lies at 40 times the Airy radius on the wing of the stellar image profile. The ratio of planetary to stellar flux lying within an Airy radius centered at Jupiter's position is approximately 10^{-5}. (This ratio varies as the third power of the ratio of telescope aperture size to wavelength.) The 10^{-5} contrast ratio would pose an impossible observing challenge for HST. The only HST instrument with any imaginable potential for adequately sampling the point-spread function at 0.25 μm wavelength is the faint object camera (FOC). However, that instrument is limited to a low count rate by virtue of real-time photoelectron counting, and from a practical viewpoint the accumulation of the 10^{10} photoelectrons per resolution element required to achieve a 10^5 signal-to-noise ratio would be impossible.

While a hypothetical planet could not be made substantially larger or more reflective than Jupiter and remain a planet, it could gain contrast against diffracted starlight if it were at a larger distance from the star. The stellar diffraction pattern falls off as the inverse third power of the angular separation, whereas the planet brightness will vary as the inverse square of this same quantity. However, long before the 10^{-5} ratio of disadvantage could be recouped, the planet's flux would fall below the detection limit of HST for a point source.

Apodization is a technique that in principle can enhance contrast for planetary detection by suppressing the wings of the stellar diffraction pattern. This normally involves masks on the first telescope focal plane and on a reimaged pupil. The first telescope focal plane is reimaged onto the detector. The apodization benefit is countered by any light scattering due to residual roughness or dust on the telescope optics. This light is not localized on the reimaged pupil; it passes the mask and forms the image wings after the pupil diffraction has been suppressed. HST has an apodization capability in the FOC, but the benefit of this feature is expected to be quite limited owing to light scattering by the residual roughness of the HST's primary and secondary mirrors.

The committee concludes that if our solar system were to be observed by a HST or SIRTF at a nearby star, Jupiter would not be directly detected. This does not preclude any attempt to utilize HST's potential advantage in spatial resolution over ground-based telescopes in a well-structured search for companions to a selected sample of stars in the solar neighborhood. Technical improvements in direct-imaging capabilities are likely in the next generation of space-borne telescopes and are discussed briefly later in this chapter (under the heading Future Observing Systems). For the present, however, only the indirect-detection methods discussed in the following section are capable of exhaustive and sensitive searches leading to valid statistical conclusions.

INDIRECT DETECTION

Since direct-detection instruments and techniques available now or probable in the near future appear to have limited potential for surveying a statistically significant population of stars, and for identifying orbiting planetary-mass objects if they exist, the three technical approaches discussed below rely on indirect planetary effects on stellar light. These include (1) image displacement due to reflex motion (astrometry), (2) Fraunhofer spectrum Doppler shift due to reflex motion (Doppler spectroscopy), and (3) modulation due to partial occultation of the star by the planet (photometry). In each case the amplitude and timing of an effect depend on the planetary mass M_P and the orbit size, given the stellar mass M_*, which can usually be determined from its spectral type and luminosity. The period, P, directly yields the semimajor axis, a, of the planetary orbit:

$$a = (M_* + M_P)^{1/3} P^{2/3},$$

where P is in years, a is in astronomical units and M_* and M_P are in units of $M_\odot$ = 1 solar mass.

Figure 5.1 shows the estimated mass distribution for observed main-sequence stars in the solar neighborhood. The median mass is about 0.3 $M_\odot$, with a 90 percent a priori probability that the mass lies between 0.1 and 0.8 $M_\odot$. The turndown below 0.2 $M_\odot$ may be a selection effect since such stars are very difficult to detect due to their faintness, but the distribution is valid for the current application. The population median 0.3 $M_\odot$ is chosen as a typical stellar mass for the calculations below.

In order to deal with a sample of a few hundred stars, a typical distance of 10 parsecs is assumed in the numerical examples that follow. For the baseline analysis, the observable effect of a single planet in a circular orbit is computed, treating M_P and a or P as free parameters. Orbital periods for planets at distances of 5 to 20 AU from a 0.3-$M_\odot$ star are 20 to 160 yr.

In reflex response to a planet in circular orbit, a star will execute synchronized, coplanar circular motion of radius $aM_P M_*^{-1}$ about the barycenter. On the plane of the sky, the apparent stellar motion that can be detected by astrometric methods will be an ellipse of (angular) semimajor axis

$$x = aM_P M_*^{-1} r^{-1}$$

where r is the distance to the star. If r is in parsecs and a is in astronomical units, then x is in arcseconds. Only the eccentricity of the apparent ellipse,

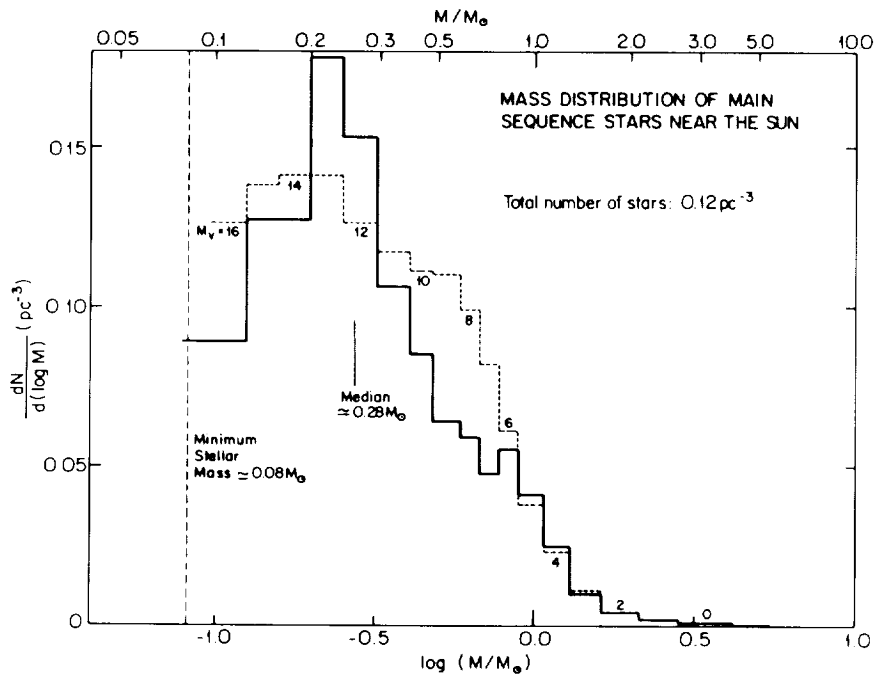


FIGURE 5.1 The mass distribution of main-sequence stars in the solar neighborhood. The dashed line represents a smoothed compilation of data from several sources by Miller and Scalo ([1979] *Astrophys. J. Suppl. Ser.* 41, 513), while the solid line is based on the more recent star counts of Wielen et al. ([1983] *I.A.U. Colloquium* No. 76). The ordinate is the number of stars per cubic parsec, per interval of log ($M/M_\odot$); the overall space density for both distributions is 0.12 stars parsec^{-3}. Also indicated are the median mass (0.25 to 0.30 $M_\odot$) and the corresponding absolute visual magnitude, M_v. The turnover in the mass distribution below $\sim$0.2 $M_\odot$ (M_v = 13) is believed to be real, but the true number of very-low-mass stars is probably underestimated because of incompleteness in the available star catalogs.

and not the amplitude x, depends on the angle between the orbit pole and the line of sight.

For a nominal Jupiter at a distance of 10 parsecs, orbiting a 0.3-$M_\odot$ star with a semimajor axis of 5 AU, x = 1.7 milliarcsec. For a nominal Uranus orbiting the same star at 1 AU, the apparent displacement would be only $\sim$15 microarcsec. This orbital displacement must be distinguished from the displacement due to stellar proper motion, which for a velocity of 10 km s^{-1} relative to the Sun amounts to 0.2 arcsec yr^{-1}. This suggests that a successful astrometric detection, even for Jupiter, requires observations over a period of time comparable with the orbital period. In any case, a determination of the orbital period combined with an estimate for M_* from the spectral type of the star is needed to estimate a, and hence to permit determination of M_P.

Ground-based astrometry using photographic plates achieves a typical accuracy of 100 milliarcsec in determining the relative positions (separations) of stars. Using photoelectric techniques at ground-based telescopes, current studies indicate that a few-milliarcsecond accuracy appears possible—which is remarkably good, considering that the center of a typical ground-based "seeing" image (1 to 3 arcsec) must be determined to 1 part per 1000 in the process. HST is expected to be able to make astrometric measurements using its fine guidance sensors (FGSs) with about this same relative accuracy.

The reflex motion of a star in response to an orbiting planet also results in a periodic Doppler shift in stellar spectral features of amplitude $\Delta\lambda/\lambda = (V/c) \sin i$, where λ is the wavelength, c is the speed of light, i is the angle between the orbit pole and the line of sight, and V is the stellar orbital speed. With G representing the gravitational constant,

$$V = G^{-1/2} M_P (M_* a)^{-1/2} = 30 M_P (M_* a)^{-1/2} \text{km s}^{-1},$$

where masses are in $M_\odot$ and a is in astronomical units. By identifying and accounting for all extraneous contributions to the line-of-sight motion, it is possible to identify small residual shifts having periodic variation indicative of an orbiting mass. Once again, knowledge of P and M_* would yield the value of the semimajor axis, a. The planetary mass could also be determined if sin i were known, but Doppler spectroscopy cannot determine it. The average value of sin i is 0.79 for randomly oriented orbits, and this value is used here to determine the sensitivity of Doppler spectroscopy for exploring the domains of a and M_P.

For a nominal Jupiter orbiting a 0.3-$M_\odot$ star with a semimajor axis of 5 AU, V = 24 m s^{-1}, and for a nominal Uranus at 1 AU, V = 2 m s^{-1}. Again, this periodically varying orbital signature must be distinguished against a background of a typical stellar radial velocity of 10 km s^{-1}, and the broadening of spectral lines due to stellar rotation (of the order of 1 km s^{-1} for a solar-type star) and a host of systematic instrumental shifts. Additional complications may arise from stellar convection and inhomogeneities in brightness across the stellar disk as, for example, in solar flares.

By careful application of Doppler spectroscopic techniques using conventional coude or echelle spectrographs, it is possible to achieve ~10 m s^{-1} accuracy in the line-of-sight motion. At a typical resolving power of $\Delta\lambda/\lambda = 5 \times 10^4$, this implies dividing and determining the centroid of a spectral line to about 1 part in 10^3. Measurements with an accuracy of this order have been reported, but this is currently restricted to observations of bright stars and requires that systematic error sources be very carefully controlled.

In principle, it is possible to search for planets by photometrically monitoring the light from stars, looking for decreases due to partial occultations during the planet's transit in front of the stellar disk. The effect is small and difficult to observe. For a nominal Jupiter orbiting a 0.3-$M_\odot$ star at 5 AU, a dip in signal amplitude of ~7 percent would be observed for a duration of ~20 hours for a diametric transit (which would occur once every 20 yr). The amplitude of the signal drop scales approximately as $R_P{}^2 M_*{}^{-1.6}$, where R_P is the planet radius in Jupiter radii, and the duration scales as $M_*{}^{0.3} a^{1/2}$ (see Appendix A). A clear identification by this technique appears unlikely.

EVALUATION OF PROPOSED INDIRECT TECHNIQUES

In this section, with the assistance of briefings by researchers expert in the principal search techniques discussed above, the committee evaluates the expected applicability of proposed and present instrumentation to the search for extrasolar planets.

Astrometric Telescope Facility

The committee has reviewed proposed plans for an astrometric telescope facility (ATF) on the Space Station that would take advantage of the smaller and more stable images available outside the atmosphere. The design of 1σ accuracy for ATF is 10^{-5} arcsec—an extraordinary goal, considering that it represents about 1 percent of the Sun's angular size subtended at 10 parsecs. Although current studies indicate that this accuracy is technically approachable, uncertainty about possible systematic errors due to offsetting of the stellar light centroid by star spots represents a potential lien against achieving the design goal of 10^{-5} arcsec.

To investigate the planetary domains accessible to an astrometric search, the committee considers the interpretation of a detected 3 displacement by an ATF-like facility, that is, the discovery of a stellar ellipse of amplitude $x = 3 \times 10^{-5}$ arcsec. It is assumed that detection is made during a program of observations of several hundred nearby stars, implying an inventory out to at least 10 parsecs or so. (The actual distance to any particular star surveyed would be accurately determined by its annual parallactic motion of amplitude $\sim 10^{-2}$ arcsec.) Adopting $r = 10$ parsecs, and recalling that a is determined from P given M_* from the spectral type, the planet mass is also determined uniquely from the observed amplitude x.

Figure 5.2 shows the locus on the (a, M_P) plane corresponding to a

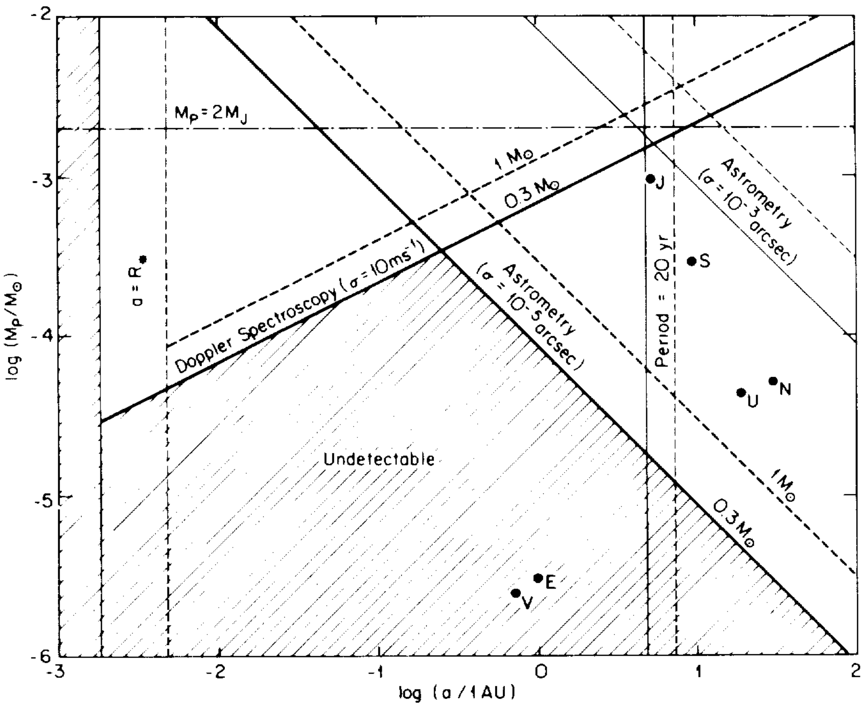


FIGURE 5.2 The discovery space for proposed astrometric and Doppler spectroscopic planetary search programs. The zone of detectability for a typical 0.3-$M_{\odot}$ star is unshaded. For astrometric searches, a 3σ detection is assumed at a distance of 10 parsecs, with $\sigma = 10^{-5}$ arcsec. Corresponding limits for an improved ground-based astrometric search ($\sigma = 10^{-3}$ arcsec) are also shown. For Doppler searches, a 3σ detection is assumed with an average value of $\sin i = 0.79$ and $\sigma = 10$ m s^{-1}. Dashed lines show the same detection limits for planets around a 1-$M_{\odot}$ star. Also shown are minimum and maximum values of the semimajor axis, set by the stellar radius and an orbital period of 20 yr, respectively.

marginally detectable (3σ) displacement for stars of 0.3 and 1 $M_{\odot}$, bounding the region of astrometrically detectable planets. It can be seen that both Jupiter and Saturn would be readily detected at a distance of 10 parsecs, provided the observations spanned a sufficiently long period of time (~20 and ~50 yr, respectively, for a 0.3-$M_{\odot}$ star). For an observing campaign with a duration of approximately 20 yr, the smallest planet detectable using this technique is of the order of 6 $M_{\oplus}$, about one-half to one-third the masses of Uranus, Neptune, and hypothetical giant planet cores.

If a star with a sufficiently large reflex motion were to be discovered, it would be possible in principle to determine both the eccentricity and inclination of the companion orbit. The intrinsically interesting case of multiple-planet systems would pose the further analytic challenge of disentangling multiple periodicities in the stellar reflex. Such an analysis, which clearly implies long observing times, would address the central question of whether or not the orbits are coplanar—as would be expected on the basis of current theories of solar system formation.

It is evident from the above discussion and from Figure 5.2 that an astrometric accuracy approaching the 1 σ design goal of the proposed space-based ATF defines an accessible discovery space that includes a relatively broad and scientifically crucial range of planetary masses. For comparison, ground-based photoelectric measurement at a site with excellent seeing (assumed $\sigma = 10^{-3}$ arcsec) could marginally detect Jupiter-mass planets in Jupiter-like orbits around a 0.3-$M_\odot$ star ~10 parsecs from the Sun—by no means a trivial accomplishment given the intense current interest in such searches and the implications of their results. But technical limitations on ground-based astrometric accuracy clearly constrain the extent to which it can address the existence, masses, and dynamics of extrasolar planetary systems, and in particular the central question of whether they do or do not resemble our own system.

Doppler Spectroscopic Planet Searches

To explore the domain of the (a, M_P) plane accessible to a Doppler spectroscopic planet search, the committee considers a search program at a 1 σ accuracy of 10 m s^{-1}. A circular orbit is assumed for simplicity, with an average value of sin i = 0.79, to estimate the radial velocity amplitude. (Note that, unlike the astrometric signature, the observed Doppler shift does not depend on the star's distance from the Earth.) As before, the committee assumes stellar masses of 0.3 and 1 $M_\odot$ and plots the locus corresponding to a 3σ detection in Figure 5.2. For a 0.3-$M_\odot$ star, Jupiter-mass and Saturn-mass planets would be detectable at the stated accuracy if their semimajor axes were respectively ~2 AU and inside 1 AU.

Although at the assumed level of precision the Doppler technique is not capable of detecting bodies as small as Uranus or Neptune at any rational orbital distance, it is apparent from Figure 5.2 that the Doppler and astrometric techniques explore complementary domains of discovery space. The former is sensitive to planets with smaller semimajor axes (and thus larger orbital velocities) whereas the latter is not—and therein, in the absence of preconceptions derived from the local example about what planetary mass-distance relationships "should" be, lies the value of the Doppler approach.

The Doppler discovery space increases with lower stellar mass, but unfortunately a reduction in mass also lowers the luminosity for main-sequence stars and so reduces the number of accessible candidates for ultrahigh-resolution Doppler spectroscopy. The ultimate sensitivity of this technique may be limited not by instrumental considerations, but rather by intrinsic noise in the spectrum due to stellar turbulence. In addition, although the value of sin i can be determined for a high signal-to-noise astrometric detection, it cannot be deduced from Doppler spectroscopic data alone. Stellar photometry sensitive to star spot rotation and spectroscopy of rotationally broadened Fraunhofer lines may give clues to the value of sin i, but it seems unlikely that these techniques will reliably reduce the associated uncertainties in M_P and in the dynamical arrangement of multiple orbits.

Several ground-based Doppler-spectroscopic planetary search programs at accuracies near or better than that assumed here are in progress. Reports noted earlier of detections of a substellar companion to HD114762 and a roughly 1-M_{Jupiter} planet orbiting γ Cephei (at about 2 AU) both utilized this technique. Neither a major technical breakthrough nor space-based system deployment is required to pursue this approach.

Photometric Planet Searches

It is clear from the Jupiter transit example (under the heading Indirect Detection) that the photometric technique must deal with events that typically may be of short duration and low frequency, and are detectable only if the observer is in or very near the planet's orbital plane. Because of the low a priori probability of meeting this requirement and thus the size of the star sample required for a minimally successful photometric planet search (see Appendix A), COMPLEX concludes that this technique as currently developed is not yet adequate to determine the statistical occurrence of planets. Detection probabilities could be somewhat enhanced by preselecting eclipsing binary stellar systems for photometric observation, presuming that possible planetary orbits would be coplanar. This enhancement would not be sufficiently robust, however, to develop statistical conclusions about the natural occurrence of planetary systems in general.

The method has several other potential drawbacks. First, if the observations were ground-based, they would be affected by variations in atmospheric transparency. Second, the intrinsic variability of stars—e.g., small variations in stellar flux caused by convection and star spots—may confuse interpretation of the data. Some observations from NASA's Solar Maximum Mission show the whole-disk variability of solar flux to be as much as 25 percent of the signal dip expected from a transit by an Earth-size planet ($\sim 10^{-4}$). Third, the interpretation of a photometric occultation event as being caused by a planet would not be unambiguous; small stellar objects such as white dwarfs and substellar objects could produce occultation effects not readily distinguishable from those due to planets.

FUTURE OBSERVING SYSTEMS

COMPLEX has also reviewed other potential approaches to the search for extrasolar planets. These may be grouped under the rubric of interferometry and generally represent attempts to improve resolution without increasing collecting area and hence expense. For the most part these systems exist at the conceptual level, with little or no engineering development to date. One specific system, Precision Optical Interferometry in Space (POINTS), is under active design study.

The committee views the design goals and anticipated capability of POINTS as representing a substantial advance over demonstrated capabilities of existing systems, or of expected direct engineering descendants of existing systems. Given its future potential, POINTS is, in the committee's opinion, worthy of continued study. The relative state of development of POINTS with respect to ATF, however, places POINTS in a candidate position for detailed follow-on studies and extended surveys after completion of an initial survey at $\sim 10^{-5}$ arcsecs with ATF-like facilities and at 10 m s^{-1} with Doppler spectroscopy.

Future application of ground-based infrared interferometry also holds the promise of detection and perhaps even imaging of extrasolar systems at high resolution. Experiments currently in progress at a wavelength of 10 μm are expected to achieve a spatial resolution of $\sim$10 milliarcsec using a baseline of $\sim$100 m. Again, this approach will initially be limited to only the brightest stars in the solar neighborhood. COMPLEX encourages further development efforts in areas such as this, because they are likely to make important contributions in the future.

The difficulty of direct imaging of planets around nearby stars was addressed earlier in this chapter. Nonetheless, such imaging is likely to be extremely important in the study of precursors to planetary systems described in the next chapter. Ultimately, the availability of large space-based telescopes of high optical quality may permit imaging of extrasolar-system planets. The contrast of a faint companion or circumstellar material against the stellar diffraction wing improves as the third power of the telescope diameter. Further, if dust contamination and residual surface roughness of the telescope optics can be suppressed, extraneous scattered light can be reduced, and innovative image tailoring by apodization can provide additionally enhanced contrast. Improved management of astronomical light within telescope systems would significantly augment the range, spatial specificity, and overall effectiveness of extrasolar planetary studies.

RECOMMENDATIONS REGARDING TECHNIQUES FOR PLANET SEARCHES

General recommendations concerning the programmatic and technical elements of a scientific strategy for detection and study of extrasolar planetary materials are presented in Chapter 8. Here the committee sets out those recommendations that specifically relate to the technical considerations discussed in this chapter:

- *Of the three different technical approaches to indirect (nonimaging) detection of remote planetary systems examined here in detail—astrometry, Doppler spectroscopy, and photometry—COMPLEX recommends that priority attention be given to the first two of these techniques.*
- *To achieve the significant improvement in astrometric accuracy needed to address seriously the detection and study of extrasolar planetary systems, space-based instrumentation, with its potential of a more than 100-fold gain in sensitivity over current ground-based instruments, will ultimately be required. COMPLEX believes that the promise of significant advancements in planetary science justifies further investigation of the technical issues involved, and recommends active development toward timely Earth-orbital deployment of such a capability. The committee further recommends that in the interim, ongoing ground-based searches be continued at their state-of-the-art accuracy, and that potential for improvement of this accuracy be investigated and implemented if technically and financially feasible.*
- *Doppler spectroscopy extends the range of potential planetary detections to planets that have a Jupiter (and possibly Saturn) mass in relatively small radius orbits and would not be detectable by astrometric techniques. This complementary technique does not require space-based facilities, and pilot investigations are already well under way. COMPLEX recommends that Doppler spectroscopic searches be continued at their state-of-the-art accuracy.*
- *Interferometric techniques from the optical to 1000 μm will eventually yield powerful tools for detecting condensed objects and mapping dust distributions around nearby stars, and for detailed imaging of relevant objects and regions. COMPLEX encourages continued development of promising Earth-based and space-based instruments and techniques of this type for follow-on detailed study of extrasolar planetary systems.*
- *With regard to imaging instruments, given their present importance in the study of preplanetary precursor systems and their ultimate potential for direct imaging of evolved planets, COMPLEX urges that the design of future telescopes incorporate diffraction control and techniques for the reduction of light scattering due to dust contamination and to residual errors in mirror figure. Appropriate technology to achieve these goals should be developed and implemented in space-based telescopes. Where appropriate and feasible, consideration should be given to improving the optical performance of existing telescope facilities envisioned for such studies.*

6
Future Investigations of Precursor Planetary Systems

The previous chapter focused on observational initiatives whose goals are the detection and enumeration of evolved extrasolar planets. The committee turns next to the question of enhancing current knowledge of possible precursors of planetary systems and protoplanetary material—e.g., evolving molecular clouds, stellar nebulae, and accretion disks—through the development of advanced laboratory and observational techniques. This development may also be critical to the study of the physical environment of candidate extrasolar-system planets when, and if, they are discovered.

The applicable techniques are *imaging*, *analytic spectroscopy*, and *polarimetry*, which will be discussed here under science objectives rather than by technique, as in the foregoing chapter on planetary searches. The reason is that clear distinctions based on technique do not exist for this area—images are taken through restrictive spectral or polarimetric filters, whereas spectra are obtained with spatial sorting or scanning. The nature of the instrumentation and subject area is such that planetary scientists will carry out those programs in association with astrophysicists. Indeed, many of the relevant projects proposed or under development spring from the astrophysics community.

In the following sections, three classes of study objects are discussed: dust, protostars, and pre-main-sequence stars. The last two were identified in Chapter 3 as the immediate precursors to evolved stellar systems, and it is in their vicinity that we can expect to observe the early stages of planet formation. The study of dust is singled out for special treatment because dust is ubiquitous and plays a unique role as a tracer of many physical processes important in stellar evolution, as well as in the formation and evolution of planetary systems.

DUST-RELATED ASTROPHYSICAL STUDIES

Recent observations, especially in the infrared, show that dust is a useful tracer of elements heavier than hydrogen and helium through the processes of star formation, evolution, and death. Dust is thought to form originally in the ejecta of evolved stars. Silicate grains have been found in the interstellar medium in regions of star formation, in pre-main-sequence stars around main-sequence stars, and in end stages of stellar evolution.

Characteristics of astrophysical grains were described in detail in Chapter 4 under Circumstellar Dust. Silicates are conspicuous to thermal infrared spectroscopy due to prominent 10- and 20-μm emission features. Silicon carbide grains emit a characteristic 10-μm spectral feature observed around some stars. Carbon grains, responsible for infrared continuum emissions, appear in all stages of stellar evolution as well as in the interstellar medium.

Dust is an important solar system constituent. Dust grains, incorporated into comets, asteroids, and planets during the formation of the solar system, represent material that was injected into the presolar nebula by earlier, evolved stellar systems and processed to various extents. Dust observed in the comae and tails of comets as it is expelled from the frozen nucleus is the probable source of the zodiacal cloud.

Dust grains of interplanetary origin can be collected in the atmosphere or in near-Earth space and studied in the laboratory to determine composition, mineralogy, and optical properties. COMPLEX has reviewed a catalog of infrared spectra taken of individual, microscopic dust particles that display a multiplicity of characteristic features relatable to their separately determined intrinsic properties and is convinced of the efficacy of techniques available for the establishment of ground truth. Dust, particularly through its thermal and optical properties, which can be calibrated, provides a useful diagnostic material that can probe astrophysical systems for knowledge connected to the solar system.

Direct knowledge of the material present in our early solar system obtained from dust-oriented studies could be compared with observations of the dusty material thought to exist in young and evolved stellar systems. A special emphasis should be placed on observations of young stellar systems where planet formation is thought to be a likely process. The dusty constituents of these systems should be carefully studied to determine how closely they resemble the material in our solar system.

Resolved observations of dust disks (or shells or clouds) can address key questions regarding the formation and evolution of planetary systems.

Determination of locations of dust shell edges and their sharpness (e.g., at resolutions of ~1 AU) will permit determination of how the dust is supplied, such as through the erosion of comets, and whether clearing of the inner edge by planets is a plausible model. For systems such as Vega, for example, one can envision future surveys utilizing a cryogenic infrared telescope or a large millimeter-submillimeter-wave dish in space to assess the frequency of Vega-like dust shells, determine gas-dust compositions by spectroscopy, and carry out direct and indirect searches for resonances and other phenomena related to radial structure. Such work links the observation and laboratory studies of dust with searches for planetary bodies discussed in the previous chapter. Studies of the dust ejected into the interstellar medium by evolved stellar systems such as supergiants, novae, and supernovae are required to trace the injection of heavier elements into young stellar systems and to determine the characteristics of the dust formed from the ejected material.

It is important to establish whether or not dust grains in the vicinity of the objects mentioned above are consistent with the hypothesis that such grains, formed in the ejecta of evolved stars, find their way into regions of star formation and young stellar systems where planets form. Observations, particularly in the infrared region of the spectrum, are needed to provide detailed information on grain formation, composition, crystallographic structure, and the relative importance of various stellar sources of interstellar dust. New observational data, combined with modeling and laboratory condensation experiments, should provide a better understanding of nucleation and grain growth. Currently there is considerable uncertainty as to how nucleation and grain growth proceed in circumstellar environments even though these processes are actually observed in classical novae. Signatures of carbon or silicate grains appear in nova spectra ~100 days following an outburst. It has been predicted that nucleation difficulties in the pure gaseous outflow from all dust-forming stars could lead to high levels of supersaturation and subsequent formation of amorphous particles, with well-mixed composition. This prediction needs to be carefully evaluated in the context of some of the isotopic anomalies seen in meteorites.

LABORATORY STUDY OF INTERSTELLAR MATERIALS

Infrared spectra of interstellar grains are complex and difficult to interpret unambiguously because there may be several constituents as well as contributions to the shape of a given spectral feature from the lattice structure of the solid material. Thus it is advisable to have laboratory ground truth analysis of candidate grain samples based on the wide range of in situ measurement techniques currently available. Such analyses will provide a vital synergistic link between dust studied with purely astronomical techniques and material that can be studied in the laboratory using sophisticated analytical techniques. Although no single presolar grain has ever been identified in meteoritic material, isotopic anomalies in these objects clearly indicate that they contain components that preserve isotopic signatures established in presolar environments. Notable among these are s-process krypton and xenon, heavy carbon, light nitrogen, and pure ^{22}Ne (presumably the product ^{22}Na decay) in carbon-rich, acid-resistant residues from carbonaceous chondrites. As discussed in Chapter 4, the preservation of ^{22}Ne is particularly significant because it implies rapid grain condensation (within a few half-lives of the synthesis of its short-lived [2.6 yr] ^{22}Na progenitor).

Future laboratory analysis of primitive solar system materials will likely lead to the identification of discrete presolar grains and not just components separable by chemical techniques. These laboratory searches should be conducted on meteorites, interplanetary dust, and samples analyzed on or directly recovered from comets by comet rendezvous or sample return spacecraft missions. Cometary samples are particularly crucial because the remoteness of comet origin and subsequent storage favors the incorporation and preservation of presolar materials.

If presolar grains can be found and shown not to have been altered significantly by solar system processes, they could be used as ground truth indicators for models of the nature, origin, and evolution of interstellar grains. The physical form and chemical and mineralogical composition of grains would provide direct tests of models of grains as either mixtures of separate silicate and graphite grains, or as composite structures with silicate cores and low atomic number (hydrogen, carbon, nitrogen, oxygen) mantles. The grain analyses could also provide information on grain modifications that occur in the interstellar medium before and during incorporation into a protoplanetary system, for example by processing in accretion shocks. The isotopic compositions of grains should contain nucleosynthetic signatures of their original stellar source. Ideally, such studies could determine the relative importance of various stellar types as sources of grains that originally formed the solar system.

If interstellar grains can be identified and clustered into groups, then critical comparative studies of the various ultraviolet, optical, and infrared signatures of interstellar dust could be made. For instance, if the collected grains showed the 2175-Å ultraviolet interstellar feature, then further laboratory tests could determine exactly what causes this ubiquitous interstellar absorption. It has been suggested that the feature is the result of collective oscillations in very small, nearly spherical graphite grains. If these existed in the collected samples, they could be imaged directly in a high-resolution electron microscope and their crystallographic and elemental composition determined. The nature of the absorption process could be tested on the $<$100-Å size scale using electron energy loss spectroscopy (EELS). The low-loss EELS peak is due to plasmon effects and, in principle, is directly related to ultraviolet spectral effects. Additional work along these lines could also yield fresh insight into the 10-μm "silicate" feature, the SiC feature, the 3.4-μm (organic?) feature, diffuse bands, and various unidentified infrared features.

An example of the potential of this field is the recent work on a single interplanetary dust sample that, in addition to other features, showed an absorption line that correlated with the 6.8-μm interstellar feature. This interstellar feature has been postulated to be related to carbonaceous matter, but in the dust grain, the feature was definitively shown to be caused by carbonates. Carbonates were identified by electron microscopy and, after dissolution of carbonates by acid treatment, the 6.8-μm feature was absent.

The search for presolar grains in the laboratory materials is limited by the state of analytical technology for small samples. Grains detected by astronomical techniques are generally in the 10^{-13} to 10^{-16}-g range, while materials analyzed in the laboratory are typically much larger. For example, a standard high-precision isotopic analysis of meteoritic oxygen requires a sample of over 10^{-3}-g mass. Recently developed ion microprobe techniques have provided isotopic analyses of carbon and hydrogen in grains in the 10^{-12}-g range. In addition to the ion probe, newly developing techniques include accelerator mass spectrometry, multicollector mass spectrometry, synchrotron x-ray microprobe analysis, ultrasensitive organic analysis, and high-resolution analytical electron microscopy. The sensitivity of techniques for analysis of extraterrestrial materials has undergone orders-of-magnitude improvement over the past two decades, but further improvements will be required to analyze presolar grains adequately. Such advances are needed for work on currently available extraterrestrial materials, and they will play a critical role when the first samples of a comet nucleus are analyzed in situ or by sample return.

In parallel with these present and future investigations of extraterrestrial materials themselves, it is extremely important to pursue laboratory studies of the processes that are thought to act on natural extrasolar grains during their life cycles from condensation to destruction or accretion. Such experiments can be conducted, using a wide variety of laboratory techniques under appropriate physical conditions (for instance, under microgravity conditions for studies of grain agglomeration), on simulant materials that appear likely to represent the spectrally inferred chemical compositions, physical states, and sizes of solid particles in stellar ejecta and in interstellar and accretion-disk environments. One example, important for the volatile inventory of planetary systems, is investigation of the condensation, chemical-organic evolution, and erosion of icy-organic grain mantles and their constituents in the last two of these environments.

PROTOSTELLAR OBSERVATIONS

A primary objective of observational and theoretical studies of protostellar and pre-main-sequence objects should be to piece together a unified picture of the temporal evolution of the physical parameters of an evolving stellar system from the onset of gravitational instability to nuclear burning on the main sequence. Such an effort is statistical in nature; many representative systems in various states of evolution (and, therefore, of different ages) must be thoroughly observed. One result of such studies would be the construction of an H-R diagram for pre-main-sequence objects.

Whereas molecular clouds and pre-main-sequence stars are common astronomical objects, an example of the intermediate, protostellar phase has yet to be definitely identified. An effective search for the elusive protostar is important. Candidates could be located by systematic imaging surveys at long wavelengths. During the protostellar collapse, cloud temperatures as low as 10 to 50 K obtain, and such systems emit predominantly in the far-infrared and submillimeter spectral regions. Systems operating at these long wavelengths can and do find cold compact objects embedded within molecular cloud cores. Initial observations, particularly by IRAS, will be followed up using a variety of more advanced facilities that are now being constructed or are planned. These include ground-based submillimeter telescopes and interferometers, proposed airborne telescopes like the Stratospheric Observatory for Infrared Astronomy (SOFIA), and space-based facilities such as SIRTF and the large deployable reflector (LDR).

Even for the nearest regions of star formation, the far-infrared and submillimeter facilities contemplated here will only resolve details as small as a few hundred to a few thousand astronomical units across. It will therefore be necessary to supplement the imaging surveys by instruments that might be expected to reveal the presence of protostellar candidates; that is, with higher-spatial-resolution imaging studies that can be conducted with very large ground-based telescopes operating in the near-infrared and thermal-infrared spectral regions. For example, the once-planned National New Technology Telescope (NNTT), with its 21-m baseline for imaging, could have resolved condensations the size of the solar system in the 10- to 20-μm spectral region for the closest regions of star formation. If large space dishes and interferometers could be optimized for this spectral region, superresolution techniques combined with interferometry could enable resolutions considerably better than this.

Such infrared studies would enable us to examine in detail the spatial morphology of the collapsing systems. It is especially important to be able to observe evolving circumstellar disks in various stages of fragmentation in order to identify sites where substellar- and planetary-mass companions may form.

Polarimetric imaging studies of star-forming cores should be useful for identifying the location of central heat sources and candidate protostars, and for mapping the spatial distribution of the dust in reflection nebulae. Spectroscopic observations of atomic and molecular lines can be used to trace the gaseous constituents and determine the spatial dependence of the physical conditions in the collapsing cloud. Dust constituents can be traced in the infrared by observing prominent solid-state emission-absorption features. It is important to recognize that imaging must be conducted at high polarimetric and spectroscopic resolutions to reveal the dynamics of and physical conditions in collapsing systems.

Fundamental information on the processes of star formation is provided by measurements of molecular line emission (e.g., carbon-oxygen, oxygen-hydrogen, and hydrogen-carbon-nitrogen) at radio wavelengths. These rotational lines are collisionally excited by hydrogen molecules and therefore can be used to determine the density and pressures within molecular clouds and the largest-scale (and thus lowest-density) circumstellar disks. Doppler shift and line-width analyses provide information on velocities, flows, turbulence, and chemical abundances. Such studies require spectral resolution in the range $\lambda/\Delta\lambda = 10^3$ to 10^6. At present, these techniques alone provide direct observational support for the rotational motion of disks. Some measurements may imply inflow within certain molecular cloud cores; hence these techniques may eventually provide unambiguous evidence for protostellar collapse and inflow. Radio measurements are currently limited by both resolution and sensitivity, especially at millimeter wavelengths where single-dish measurements are common.

OBSERVATION OF PRE-MAIN-SEQUENCE STARS

In contrast to protostars, many pre-main-sequence stars have been identified. They exist in regions of dense obscuration, a factor that favors longer-wavelength studies because of lower light extinction. These objects exhibit a range of identified or inferred phenomena—mass outflows, T Tauri outbursts, accretion disks, Herbig-Haro objects, and expanding shocked shells—and they are accessible to the full suite of imaging, polarimetric, and spectroscopic studies. It is particularly important to address the key questions related to planetary evolution in studies of such objects.

Studies of the immediate post-protostellar collapse phase, when the central regions of the collapsing system are heating up, can be fruitfully conducted at the highest spatial resolutions using the near-infrared and even optical techniques on the largest telescopes. Spatial scales much smaller than 0.1 arcsec, which represents a 15-AU resolution in the closest regions of star formation, could be probed using interferometers in space. General imaging and photometric studies combined with polarimetry can reveal the spatial morphologies of collapsed disk systems and the outflows associated with pre-main-sequence stages of stellar evolution. Observations of the smallest structures in evolving pre-main-sequence stars will require long-baseline interferometry, perhaps from orbiting telescopes where the deleterious effects of atmospheric "seeing" can be avoided and where orbital precession can be used to provide complete spatial coverage.

Using the techniques of infrared spectroscopy, imaging spectroscopy, and polarimetric studies, we can expect to obtain new structural and dynamical information about pre-main-sequence stars. At the highest spectral resolution ($\lambda/\Delta\lambda = \sim 10^4$ to 10^6), spectroscopic studies will provide information about abundances and dynamical motions. The spatial resolutions that would be available with LDR, and with ground-based facilities as competent as the once-planned NNTT, would enable studies of turbulence during early stages of stellar evolution. Observations of the differential rotation of nebulae can be used to address the angular momentum transfer processes.

Spectroscopic imaging studies of the atomic and molecular lines and dust emission features can reveal important information about the ratio of gas to dust and the chemical abundances of these species as a function of position in a nebular system. There is ample evidence from studies of our own solar system that spatial chemical fractionation by condensation occurred during the nebular phase. An especially important line of investigation will be to determine the chemical evolution of nebular systems as a function of time. Spectroscopic measurements in the near-infrared, far-infrared, and microwave spectral regions can reveal how complex organic molecules that are created in molecular cloud cores are incorporated into protostellar condensation, and how these molecules evolve as the process of stellar collapse and condensation progresses. Similar studies of the fate of the elements themselves (such as the biogenically important carbon, nitrogen, and oxygen) can also be pursued. Spectral line studies will additionally reveal the nature of the outflows in pre-main-sequence stars and can be expected to shed light on how the outflows interact with the surrounding molecular cloud and circumstellar environment.

The energetic phenomena associated with pre-main-sequence stars involve large degrees of ionization and powerful magnetic fields. Radio measurements are ideally suited to providing information on emission mechanisms, electron temperatures and populations, and magnetic field strengths and configurations. Energetic plasma phenomena vary rapidly in time, and so continuous radio monitoring can in principle follow the details of accretion-disk inflow, star spot evolution, jet acceleration, and so on.

At present, these measurements are carried out at instruments such as the sensitive, high-resolution (subarcsecond) Very Large Array (VLA).

SPECTROSCOPIC STUDIES OF EXTRASOLAR-SYSTEM PLANETS AND PREPLANETARY MATERIALS

Beyond its use for detection and fundamental characterization by the techniques discussed in Chapter 5, spectroscopy is a potential tool to elucidate physical and chemical environments of planetary and precursor systems. For reasons discussed under the heading Direct Detection, this is beyond the capabilities of telescopes or other systems currently in progress or planned, although there appears to be no fundamental reason why improved systems with the needed capabilities cannot and will not be built. Application of analytic spectroscopy to physical and compositional studies will require telescopes of sufficient spatial resolution to separate emitted or reflected light from the planetary object and circumstellar material from that of the primary star.

If such systems were trained on extrasolar planets, spectroscopy could provide the means for ascertaining the presence of an atmosphere, determining or constraining its composition, and establishing a rough temperature-pressure structure. Such measurements, combined with studies of metal content of the parent star, could be diagnostic of the processes by which the secondary object formed. In addition, emission spectra from substellar objects (which should be obtainable in the infrared, for example, from SIRTF) give information on the composition, temperature, and surface gravity of these objects. Basic questions to be addressed might concern the frequency of substellar objects and of planets with atmospheres and among those planets the determination of the frequency of occurrence of major atmospheric types, such as those primarily of hydrogen or carbon dioxide, if the solar system is used as a guide, or perhaps even other constituents. The detection of molecular oxygen as a major atmospheric constituent may even suggest the possibility of the existence of Earth-like biological activity.

Prerequisites for application of spectroscopic techniques include the existence of imaging or interferometric systems with sufficient spatial resolution and a large enough aperture to collect the required number of photons. Thus application of spectroscopy to physical studies will follow the development of light-collecting and sorting facilities in a natural way. The development of technology for spectroscopy of extrasolar-system planets is certain to be a formidable challenge, requiring extensive laboratory development of narrow-band spectrometers for the visible and infrared that are analogous to the radio-frequency spectrometers used with radio telescope arrays. For example, development of heterodyne spectrometers for the far infrared has progressed markedly in recent years. Future extensions to shorter wavelengths should find wide applicability to the spectroscopy of protoplanetary nebulae and other systems whose content is primarily in molecular form. It is likely that new molecular and atomic laboratory data will be necessary to support analysis of these observations. Thus, if NASA is to be ready for further studies when other planetary systems are located, it must begin to refine spectroscopic capabilities for that purpose.

SUMMARY RECOMMENDATIONS

- *The committee strongly supports continuation and expansion of observing programs directed toward investigation of preplanetary environments and precursor materials, and for development of enhanced imaging, spectroscopic, and polarimetric instrument capabilities relevant to such programs.*

In addition, COMPLEX makes three more specific recommendations with respect to the dust-related research discussed above, and later in Chapter 7.

- *Observational studies, collection programs and techniques, and laboratory investigations focused on extraterrestrial dust should be continued and refined.*
- *Within this general area of study, the technology of laboratory analysis of small extraterrestrial samples should be developed to the point that submicron grains carrying "exotic" isotopic signatures suggestive of presolar origin can be individually identified and analyzed.*
- *Finally, active encouragement should be given to theoretical, observational, and laboratory simulation studies of the condensation and chemical nature of dust grains in circumstellar environments, of grain interactions with and survival in the interstellar medium and during infall into accretion disks, and of the chemical and physical properties of icy-organic grain mantles.*

7
Relation to Other Astronomical and Astrophysical Studies

The effort to detect objects near other stars will open up new observational capabilities that will have a decisive influence on the development of knowledge in related astrophysical research. Some of the fundamental questions of astronomy, such as the formation and evolution of stars and the formation, dynamics, and evolution of galaxies, require for their clarification much more accurate data than are available at present. In the following sections a few examples are given that connect the study of extrasolar planetary materials with other areas of investigation.

FUNDAMENTAL PROPERTIES OF STARS IN THE SOLAR NEIGHBORHOOD

The search for planetary systems will naturally involve an intensive study of the stars in the solar neighborhood, out to a distance of perhaps 100 parsecs. The fundamental properties of most of these objects, particularly those of low mass relative to the Sun, are not accurately known; they are crucially important, however, both for comparison with theories of star formation and stellar evolution and for establishment of the extragalactic distance scale. These properties include distances, masses, radii, luminosities, the mass spectrum, and the binary or multiple system frequency. An astrometric telescope with precise measurement capability would have a substantial impact on the accurate determination of distances by direct trigonometric parallax and therefore, indirectly, on most of the other quantities.

In the area of star formation, two important pieces of information would influence ideas on its relation to planet formation. First, what is the nature of the stellar mass function at the low-mass end? It is thought to peak at about 0.25 $M_{\odot}$; the number of stars per unit volume per unit mass interval declines for lower masses (see Figure 5.1 in Chapter 5). Observational data in this range are very uncertain, however, as there are large selection effects. Do substellar companions and planets lie on a continuum extending to the lowest masses, or are there fundamental discontinuities in the mass spectrum?

Second, what is the statistical frequency of binary systems of various mass ratios? Is the frequency of binary or multiple systems different among systems in which the primary is at the low-mass end of the main sequence? How common are substellar companions vis-a-vis planetary companions? Does the number of stars per unit mass interval really drop off sharply below 0.25 $M_{\odot}$? Such data should provide clues regarding the fundamental question of whether low-mass stars and substellar objects form by the same process as do giant planets. Are most substellar companions products of binary star formation, perhaps due to high angular momentum? Are most planets products of the accretion of dust?

In addition, information about main-sequence stars in the solar-mass range is extremely important, because the evolutionary lifetime is comparable to or shorter than the age of the galaxy. Theoretical models of stars and stellar evolution are compared with observed properties of stars, including their masses, luminosities, and surface temperatures. The purpose of the comparison is to obtain fundamental knowledge about the energy source, age, composition, and detailed structure of stars at various stages of their evolution. At present, masses of only a small number of stars have been determined with reasonable accuracy. Astrometric programs and radial velocity programs to search for extrasolar planets would undoubtedly yield a number of cases in which stellar masses could be measurable by the classical method of determining the orbits of a binary system. Furthermore, the accuracy of existing measurements of mass as well as binary statistics could be greatly improved. It is interesting to contemplate the detection of a black hole by the same astrometric technique as that used for planetary detection. It would also be desirable to have accurate luminosities of nearby stars; for these, distances are the greatest uncertainty.

An astrometric facility with the accuracy needed for planet searches could result in dramatic improvements in fundamental distance measurements for nearby stars. A further consequence of improved distance measurements would be increased accuracy of observational determinations of stellar radii, for example, by fitting infrared measurements to model stellar atmospheres. Thereby the accuracy of the empirical mass-luminosity relation and the empirical temperature-luminosity relation on the main sequence could be substantially improved and the observations extended to a much greater volume of space. Even with a resolution of 10^{-4} arcsec, stars within 100 parsecs could have their distances measured to 1 percent accuracy, a great improvement over the measurement capability of the European Space Agency's Hipparchos satellite.

As one example, the luminosities of subdwarf stars in the solar neighborhood could be much more accurately determined. These objects lie below the ordinary main sequence in the H-R diagram, have low heavy-element abundances compared with those of the Sun, and are thought to be among the oldest stars in the galaxy. The determination of their precise location in the H-R diagram is an essential step in determining the absolute location of the main sequence of globular clusters. The H-R diagrams of these clusters, in connection with the theory of stellar evolution, are analyzed to give the age of the galaxy, one of the most important parameters in cosmological theories.

PROPERTIES OF STAR-FORMING REGIONS

The investigation of star-forming regions for evidence of protostellar objects, protoplanets, and nebular disks is likely to yield additional information on some of the fundamental questions about star formation. Examples include observational identification of protostars actually in the gravitational free-fall collapse phase; observation of bipolar outflow and its relation to star and planet formation; clarification of mechanisms for binary star formation and the relationship between the formation of binaries and disks; the role of magnetic fields, turbulence, and thermal pressure in initiation of star formation and during the collapse phase of protostellar condensations in molecular clouds; the lifetime of nebular disks, which determines the time available to form planets; and detailed studies of the radiation coming from protostars at all stages of their evolution.

Questions related to the dust in star-forming regions are closely connected to questions about planet formation. Does the presence of circumstellar dust automatically imply accretionary growth toward subplanetary masses? Is the dust usually blown away by stellar winds before it has a chance to accrete, or can the dust just stay in orbit around the star without forming planets? Has the dust observed in young stellar objects been injected into the interstellar medium by evolved stars, or does it nucleate and grow during the collapse phase? Do the infrared excesses recently discovered by IRAS around seemingly ordinary main-sequence stars like Vega imply that dust has already aggregated, perhaps yielding subplanetary masses resembling our asteroid belt, or are these places where planet formation has failed? In the latter case, Poynting-Robertson drag effects would result in dissipation of the dust disk in $\sim 10^6$ to 10^7 yr, and some mechanism for replenishment would be required. Is there evidence for the formation of halos of dust and gas, where cometary material could be accreting, in disklike young stellar objects?

More generally, what are the statistics of the presence, the structure, and the evolutionary time scales for such disklike objects, and how might high-precision astronomical measurements from the ground and in space best provide the basis for diagnosing these statistics? Is the formation of disks with sizes comparable to the solar system and masses comparable to the minimum mass solar nebula a common outcome of the star-formation process? A further question involves the star-formation rate in molecular clouds, which is believed to be limited by turbulence within the cloud. Bipolar outflows may provide the energy source that sustains the turbulence; in such a case, star formation would be self-regulated. Or is the magnetic field the main mechanism for limiting star formation?

Studies of the interaction between the outflows and the surrounding molecular cloud might be expected to provide crucial tests for any one of these hypotheses. With orbiting telescopes such as HST and SIRTF, it should become possible to study several hundred systems with ages between 10^5 and 10^6 yr to determine such properties as rotation, flattening, ringlike structures, and the role of winds in disk evolution and dissipation.

THEORY OF STELLAR EVOLUTION

The existence of planets, particularly massive ones, could generate observable effects in certain advanced phases of stellar evolution. When a normal star evolves from the main sequence to the red giant region, its radius expands by several hundred times. A large planet in orbit would be encountered by the expanding star and would begin to spiral in through the envelope of the giant as a consequence of frictional drag. The planet could accrete mass from the envelope, and at the same time the stellar wind from the giant would begin the process of ejection of the envelope.

The possible outcomes of this evolution range from complete evaporation of the planet to its spiraling down into the interior of the giant. If the combination of circumstances were favorable, the planet, having gained mass by accretion and having lost angular momentum by frictional coupling with the envelope, would end in a short-period orbit about the dense core of the giant. At the same time the low-density envelope of the giant would have been expelled.

This type of system, composed of an object of about 0.1 $M_\odot$ (the augmented planet) in orbit about a white dwarf (the core of the red giant), has characteristics similar to those of the so-called cataclysmic variables, which are deduced to be binary systems whose outbursts in light can be interpreted as a consequence of the transfer of mass from one component to the other. The buildup from planetary mass to stellar mass in this process could represent a new and unusual mode of star formation.

Many fundamental questions regarding the processes of stellar evolution can be addressed through studies of dust. Dust is believed to carry the heavy-element products of stellar evolution through the cycle from the interstellar medium (ISM) into star-forming systems and then back into the ISM in the ejecta of evolved stellar systems. For example, silicate and carbon grains have now been observed in the ejecta of supergiants, novae, and planetary nebulae, in the material producing the general interstellar extinction, in molecular cloud cores, in pre-main-sequence objects embedded in these clouds, and in the comae and tails of comets in the solar system. Do grains that form in the ejecta of evolved systems survive to be incorporated into young stellar systems? What is the rate of destruction by interstellar shocks? Do meteoritic inclusions with anomalous isotopic compositions represent condensations formed in the ejecta of such transient energetic nucleosynthetic events as nova and supernova explosions? Do cometary dust grains and the zodiacal dust grains resemble the dust grains in the ISM, in circumstellar shells of evolved stars, and in regions of star formation? How do grains participate in such processes as the onset of stellar collapse, planetary accretion, and the coupling of outflow energy to molecular clouds?

Studies of dust in connection with the formation of planetary systems would help to illuminate elements of the larger processes that involve the generation of heavy elements in stars, the ejection of this material into the ISM and its incorporation into grains, and the subsequent cycling of the material into new stars. Because of their general astrophysical importance and their particular relevance to investigation of precursor planetary systems, the committee sets out specific recommendations concerning dust-related studies (see Chapter 6).

DISKS IN ASTROPHYSICS

The observation and theoretical interpretation of disk structures are critical elements in the study of extrasolar planetary materials. The improved understanding of disks that could be obtained by intensive investigation of relatively nearby systems, of the T-Tauri-type and others, could influence studies in other areas of astrophysics. Disk systems play an important role in stellar binary systems in which mass is transferred between the stars. Examples are cataclysmic variables and novae. Other systems whose theoretical interpretation involves disks include the unusual galactic object SS433 and quasars. One of the crucial elements in all discussions of disks is the mechanism for mass and angular momentum transport; some kind of turbulence is indicated, but its origin is completely unclear in most systems. In the specific case of the presolar nebula, there is a theoretically suggested physical basis for it.

It is evident that theoretical studies of the "solar" nebula could result in considerable input to studies of other disk systems, including consideration of such questions as evolutionary time scale, the physics of turbulence, the stability of the disk to both thermal and dynamical effects, and tidal effects. A particular area of interest appears to be the interactions between disk systems and the planets that may form within them. Recent studies of ring systems in our solar system have revealed that density waves produced by orbital resonances profoundly affect any structure in, for example, the Saturn system. This phenomenon could have wide application in other systems, such as disks in binaries.

As a further issue, the existence of jets in extragalactic objects has been linked to the presence of disks. In this connection the jets observed near pre-main-sequence objects as part of the complex of observations known as bipolar flows are also thought to be influenced by the presence of a disk. Is there a common mechanism involved, which can operate on widely different scales? Improved theoretical and observational understanding of both the generation and collimation of bipolar outflows near young stars could therefore have wider implications in the study of a vast range of other astrophysical phenomena.

PROBLEMS IN GALACTIC STRUCTURE

Progress on a number of problems in galactic structure, such as the distribution of mass in the galaxy, the metal content of the various components of the galaxy, and the origin of the globular clusters with respect to the origin of the galaxy itself, is critically dependent on knowledge of stellar distances and motions. One of the instruments proposed in this report, a large astrometric telescope with the capability of measuring displacements small enough for the detection of planets, could also be used fruitfully in the investigation of such problems.

For example, through measurement of their trigonometric parallaxes, the distances of Cepheid variables could be measured directly for the first time. These objects, through the relation between their oscillation periods and absolute luminosity, provide a fundamental "standard candle" for establishment of the distance scale in the local group of galaxies and in the universe as a whole. The accurate measurement of the distances to even a few of the nearer Cepheids would provide the long-sought after "zero point" for the period-luminosity relation. An astrometric telescope with 10^{-4}-arcsec precision could determine the distances to at least 20 Cepheids within 5 percent. Distances could also be measured for other interesting classes of stars such as the RR Lyrae variables.

Accurate distances to these and other types of stars, combined with measurements of stellar motions at different distances from the galactic center, would contribute to the determination of the density distribution of stars and the galactic gravitational potential in both the spheroidal and disk components of the galaxy. The rotation curve of the galaxy within a few kiloparsecs of the Sun could be more accurately determined. An astrometric telescope, if used at near-infrared wavelengths, could measure the motions of stars and other features in the vicinity of the galactic center, improving our knowledge of that fascinating region and testing the hypothesis of the existence there of a black hole.

The study of globular clusters is critical for the understanding of galactic structure. The determination of their space motion with respect to the galactic center would be possible for the first time with precise astrometric observations. For instance, a cluster at a distance of 1 kiloparsec from the Sun with a transverse motion of 100 km sec^{-1} would show a displacement on the sky of 0.2 arcsec in 10 yr. These measurements would lead to the distances of the globular clusters, would provide a determination of the galactic gravitational potential at large distances from the Sun, and would allow the mass-to-light ratio of the galaxy to be deduced. This observed quantity is crucial for the establishment of the properties of the cold, dark component of the galactic mass. One would also be interested in measuring globular clusters of various metal contents, in order to determine if there is a correlation of metal content with orbit. In addition, the membership of a globular cluster could be determined, and the individual orbits of stars in the cluster would be measurable, giving an indication of the dynamical evolution of the cluster.

Some measurements of the relative motions of other nearby galaxies with respect to our own galaxy would also be possible—the proper motions of the Large and Small Magellanic Clouds, for example, correspond to displacements of 4 milliarcsec in 10 yr. With these results one would be able to determine the dynamics and masses of the systems, deduce the history of past interactions, determine the direction of the angular momentum, and possibly probe the gravitational potential of our galaxy at very large distances.

IMPLICATIONS REGARDING THE SEARCH FOR LIFE

This report does not deal with various attempts to search for extraterrestrial intelligence, such as radio telescope "listening" surveys, as a mode of searching for planetary material around other stars. Yet the two endeavors have an obvious connection. For example, if artificial radio signals were detected, it would presumably indicate the evolution of planetary systems, as well as life, somewhere in the galaxy; the habitable planets would probably, but not necessarily, be in the star system from which the radio signals emerged.

Perhaps more realistic than an imminent discovery of extraterrestrial life is a scenario in which objects in the planetary-mass range are detected without any concomitant evidence of life. Detector systems are evolving with such rapidity that we can anticipate positive identification of extrasolar planets, if they exist, within a decade or so of the initiation of a comprehensive search. Current theories lead us to suspect that other planetary systems, habitable planets, and perhaps even life forms are likely, but as yet we have no direct confirmation that even a single extrasolar planet exists. An actual discovery of planets would give the search for extraterrestrial life more impetus and a somewhat firmer scientific basis than it has today.

If planets were detected, an immense and immediate popular and scientific interest would be generated. People would want to know if the planets are Earth-like, or habitable. A spectroscopic search for oxygen or radio listening would suddenly be given specific targets. Viewed in this way, a successful search for evolved planetary systems around other stars can be seen as stimulating the search for life elsewhere in the universe.

8

Recommendations

RECOMMENDED PROGRAM

In previous reports, COMPLEX and the Space Studies Board have set out strategies for exploration of the inner planets, outer planets, and primitive bodies within our solar system. These explorations, by spacecraft, remote sensing from Earth, and laboratory studies of meteorites, lunar samples, and interplanetary dust, are leading to enormous strides in our knowledge of the physical and chemical state of the solar system, and to a certain extent also of conditions and processes pertaining to its origin and early evolution. But in a broader sense, the depth of our understanding of solar system origin, the mechanisms that produced its regularities of composition and structure, and the degree to which these processes and conditions were unique to our system among the myriad of solar-type stars in the galaxy, have been fundamentally limited by access to only one solar system for detailed observational study.

While it is clear from the discussions in Chapter 3 of this report that much theoretical work remains to be done, it is equally clear that models of solar system origin have matured to the point where they make specific predictions about processes and configurations we would expect to observe in other stellar environments that are evolving toward our type of solar system (e.g., dissipative accretion disks, coplanar planetary orbits, and the like). At the same time, several of the observational techniques discussed in Chapter 5 have also matured to a level where the detection and first-order characterization of extrasolar planetary systems appear to be within observational reach with current technology, or with feasible technological advances.

This parallel development of the solar system data base, theoretical interpretation, and capabilities for observation beyond the solar system potentially marks the beginning of an era of exploration. Many fundamental and unanswered questions now can be both meaningfully defined and observationally addressed, not only concerning just the existence of other planetary systems, but also the extent to which our type of planetary configuration is either rare or unique—or, alternatively, a frequent result of natural evolution by similar processes from common initial stellar conditions. The profound scientific and intellectual significance of even partial answers to such questions is evident, and *COMPLEX accordingly recommends that the NASA Office of Space Science and Applications (OSSA) initiate a formal program of investigation of extrasolar planetary systems and materials as a logical and scientifically necessary extension of the exploration of our own solar system. This new and intrinsically interdisciplinary program must incorporate diverse elements of investigation that have traditionally been carried out within the separate disciplines of planetary science, astrophysics, and astronomy. Its implementation will therefore require scientific and programmatic coordination across these traditional boundaries.*

The inherent scientific breadth and importance of this effort extend far beyond the simple "existence theorem" for extrasolar planets to include study of statistical distributions among star classes, of the structure and dynamical behavior of evolved planetary systems, and of circumstellar materials and processes that may precede or postdate planetary formation. *Therefore COMPLEX further notes that full implementation of the search components of this program will necessarily involve examinations of a statistically significant number of candidate stars. These must be carried out by observations utilizing a variety of techniques and extending over periods of time sufficient to demonstrate, with reasonable probability, the presence or absence of precursor or evolved planetary systems, and to characterize such systems if they are found.*

SCIENTIFIC OBJECTIVES

From the discussions in Chapters 5 and 6 of this report, it is clear that there are multiple approaches and a wide range of observational techniques applicable to the recommended program. It is also evident that the measurement capabilities required to achieve an appropriate level of scientific return from the program generally exceed those of existing ground-based astronomical instruments and of orbital telescopes currently in operation or under development. Because of the extreme observational challenge posed by the needed measurements, new astronomical instrumentation developed specifically for study of extrasolar-system materials must include Earth-orbiting telescopes and ancillary analytic equipment. The measurement requirements imposed on these instruments are, however, comparable to those desired for many areas of stellar and galactic astronomy. Thus the technological imperatives of the program can best be addressed and implemented by coordination in instrumental design and operation, and by sharing of available resources and facilities, with other related disciplines. It is important to note, however, that the search programs required for detection and detailed study of extrasolar planetary materials impose important constraints on the commitment and duration of observing time, and could be severely compromised if truncated by overassignment of other scientific tasks to crucial instruments or facilities.

It is also possible that additional initiatives may be warranted, based on new and currently unforeseen advances in any of several technologies, which may lead to significant breakthroughs in our ability to achieve the goals stated in Chapter 2. Promising initiatives of this kind should be encouraged and supported through existing instrument-development programs. Finally, the committee emphasizes the necessity of maintaining strong laboratory and theoretical programs, complementary to the observational enterprises, as essential elements of the overall program.

Based on the review in Chapter 4 of the current status of research in investigation of extrasolar planetary materials, and the evaluation in Chapter 5 of the technological and observational initiatives needed to implement the recommended program, *COMPLEX has developed the following primary scientific objectives that in the opinion of the committee are attainable during the next decade:*

1. *To search for evolved extrasolar planets and planetary systems.* This objective requires systematic observational planet searches that encompass the widest feasible domain of the planetary mass versus semimajor axis exploration space displayed in Figure 5.2. Specifically, to

a. Initiate and carry forward an astrometric observational survey program, extending for at least a decade and designed to track the reflex motion of 100 or more stars in the solar neighborhood ($r \leq 10$ parsecs) with the precision specified in the Measurement Requirements section below; and

b. Obtain and interpret a record of Doppler shifts in stellar spectral features due to reflex motion, at or above the current measurement accuracy of ~ 10 m s^{-1}, in a survey of the duration and extent specified for the astrometric survey.

2. *To characterize precursors of planetary systems.* This objective requires continuation and augmentation of present observational studies of known young stellar systems, and of the physical properties of circumstellar dust assemblages as precursors to and products of planetary systems, on a variety of spatial and spectral resolution scales. It also requires a survey of a statistically meaningful number of stars of varied masses and type to determine the frequencies and properties of other systems of this kind.

3. *To characterize precursor planetary materials.* This objective requires continued and refined investigations of links between interstellar-circumstellar dust and isotopically "exotic" grains in solar system materials such as primitive meteorites, interplanetary dust, and comets. Important specific objectives of this effort include careful collection—in the stratosphere and on Earth-orbiting facilities—and appropriate curation of rare interplanetary asteroidal-cometary dust particles (IDPs); laboratory identification and analysis of presolar dust grains preserved in these meteoritic and IDP materials; and laboratory simulation of the formation and physical and chemical processing of interstellar grains in preplanetary and planet-forming environments.

4. *To improve the capability of theoretical models and computer experiments.* This is necessary to make specific predictions regarding the observational properties of planetary systems at all stages of their evolution, and to further develop models to aid in the interpretation of existing data. In particular, it would help encourage development of improved numerical techniques such as hydrodynamical computation in two and three dimensions, coupled with radiative transfer and chemical modeling of such systems.

MEASUREMENT REQUIREMENTS

Two types of astronomical observations are specifically encouraged in this report: searches for evolved planets, and physical studies of the precursors and products of planetary systems. The committee has additionally emphasized the importance of laboratory analyses of extrasolar matter preserved in solar system materials, and of experiments to simulate the physical and chemical processing of dust grains in astronomical environments. *COMPLEX outlines below the set of measurement requirements that in its opinion are technologically attainable within the next 10 yr and are needed to achieve the stated scientific objectives of the decadal strategy.*

1. For the planet searches, the committee finds that the instrumental performances adopted in Chapter 5 are practical limits in two senses. First, $\sigma = 10^{-5}$ arcsec for astrometry and $\sigma = 10$ m s^{-1} for Doppler spectroscopy appear to be technological limits imposed by current abilities to design, construct, calibrate, and maintain astronomical instrumentation. Second,

at these levels of sensitivity both techniques will record signals originating in stellar stochastic variability—turbulence, star spots, and the like—that will introduce systematic noise effects interfering with interpretation of the data as they relate to stellar reflex motion. (It should be noted, however, that while such signals may be unwelcome "noise" for detection of reflex effects, they will contribute sensitively to knowledge about these kinds of stellar variability and in this sense will be valuable.) In the absence of better understanding of the relevant aspects of stellar activity, more sensitive measurements than those recommended below are thus not currently useful for planetary searches by these techniques. *Therefore the committee recommends, for astrometric and radial velocity planet searches:*

a. *That the design goal for relative astrometric accuracy be σ = 10 microarcsec.* If our solar system could be observed for the required times from 10 parsecs with this accuracy, the ellipse amplitude for Jupiter would be 50σ, and Uranus and Neptune would generate reflex motions of amplitude 8σ and 16σ, respectively. It should be clearly understood, however, that a 10-yr astrometric observing program, while tracking Jupiter for nearly a complete cycle, would follow Neptune and Uranus for only 5 to 10 percent of their orbital periods and would thus require longer periods of observation for unambiguous detection and period determination.

b. *That Doppler spectroscopic observations be carried out at a sensitivity of 10 m s^{-1} or better for the velocity of the orbital reflex.* Such accuracy would detect a Jupiter-mass object within the orbital radius of Venus around a solar-mass star, or at the orbital radius of Jupiter around a 0.1 $M_{\odot}$ star. Inclusion of Doppler measurements in the planet search strategy opens a unique region of the discovery space diagram, one sensitive to relatively low-mass planets in close orbits (see Figure 5.2).

c. *That both the astrometry and Doppler spectroscopy search programs be maintained continuously for at least a decade.* An extended period of observation is of the utmost importance to a successful outcome.

2. *For comprehensive physical studies of precursor systems and planetary products, COMPLEX finds that observations from x-ray to radio wavelengths are required on a variety of spatial and spectral scales ranging from moderate to very high resolution.* For example, searches for protostellar candidates may initially proceed at moderate spatial resolution (1 to 5 arcsec), but studies of the structure and dynamics of evolving protostars will require the highest spatial resolution we can muster (<0.1 arcsec). While many spectroscopic observations of dust in precursor systems can be conducted at spectral resolutions of $\lambda/\Delta\lambda \sim 10^2$, dynamical and kinematic studies of collapsing objects and bipolar outflows, analyses of chemical abundances, and determinations of excitation conditions require very high spectral resolution ($\lambda/\Delta\lambda \sim 10^4$ to 10^6). *To investigate links between interstellar-circumstellar*

and solar system dust grains, the committee urges refinement of both observational and analytical techniques. Infrared observations of novae and pre-main-sequence systems with subarcsecond spatial resolution and spectral resolutions of $\lambda/\Delta\lambda > 10^2$ are needed to provide information on circumstellar dust condensation and evaporation sequences that can be directly compared with condensation models and relevant solar system materials. Ultraviolet and infrared observations of comets require improvement to an extent sufficient to provide meaningful comparisons of the spectral signatures of cometary and extrasolar dust: at a minimum, this should include a sensitive search for the 2175-Å feature, high-resolution spectroscopy of silicate features, and searches for all major infrared signatures seen in interstellar dust.

3. *For laboratory studies of dust, COMPLEX recommends continued support and augmentation of current efforts to refine laboratory techniques and instrumentation to the point where micron to submicron extrasolar grains preserved in primitive solar system materials can be identified, isolated, and analyzed.* For micron-sized grains, chemical and isotopic requirements are determination of elemental compositions for major elements, including the low atomic number elements carbon and nitrogen; measurement of elements heavier than silicon at the 100-ppm level; isotopic analyses of carbon, hydrogen, oxygen, and nitrogen, when they are major elements, to an accuracy of 1 percent; and similar or better isotopic accuracies for several heavier elements, including magnesium, silicon, and calcium. Laboratory ultraviolet, optical, and infrared studies of such grains are also needed, with resolution and sensitivity comparable to those obtained for observation of dust in astronomical environments. *The committee further urges that active encouragement be given to laboratory simulation and theoretical studies of the astronomical dust cycle,* including the condensation and chemical nature of dust grains in circumstellar environments; the formation, evolution, and chemical and physical properties of icy-organic grain mantles; and grain and mantle processing in the interstellar medium and during infall into accretion disks and incorporation into subplanetary objects.

STRATEGY AND IMPLEMENTATION

The initial decadal science objectives of the recommended program, set out earlier in this chapter, specifically emphasize the new development or augmentation of observational, laboratory, and theoretical capabilities, and subsequent survey modes of astronomical search and study. This emphasis clearly defines a reconnaissance phase of exploration strategy. Requirements for measurement accuracy are necessarily severe because of the extreme observational and analytic challenges posed by the nature of the targets. In previously recommended strategies for the study of our own planetary system, initial reconnaissance has aimed for rapid first-order characterization, not necessarily at high levels of precision. It has been followed relatively quickly by a transition to more lengthy stages of detailed exploration and intensive study, implemented by specific spacecraft missions, designed to address well-defined questions, to known individual targets.

For investigation of extrasolar planetary materials the duration of the reconnaissance phase is more difficult to assess, although it is likely to be long for planet searches in particular. We may anticipate that it will result, on several different time scales, in the definition of targets—in space, in the laboratory, and in theoretical development—for more detailed exploration and intensive study. It is important, during this reconnaissance, to develop increasingly sensitive and sophisticated instruments and techniques for follow-on observational and laboratory studies. This would help prepare technologically to exploit these targets once they are identified.

In this context, COMPLEX offers the following recommendations and advice on exploration strategy and programmatic implementation to the Office of Space Science and Applications (OSSA):

1. *The committee recommends development, fabrication, installation, and long-term support of an astrometric observational facility that meets the astrometric measurement requirements specified above in Measurement Requirements.*

2. *The committee recommends that planetary search programs utilizing astrometry and Doppler spectroscopy at their current state-of-the-art sensitivities, be established and supported for a minimum observational period of 10 yr following their initiation. The committee further recommends that in the interim, ongoing ground-based searches be continued at their present best accuracies, and that the potential for improvement of these accuracies be investigated and implemented if technically and financially feasible.*

3. The committee recognizes the profound importance of understanding the mechanisms of star formation in the galaxy for theories of the formation and evolution of our own planetary system. To develop such knowledge requires comprehensive investigation of possible precursors and products of extrasolar planetary systems. This in turn requires a broad range of coordinated observational, laboratory, and theoretical studies as enumerated above in the Scientific Objectives and Measurement Requirements sections of this chapter. *The committee therefore recommends that current studies in these areas be continued and augmented and that new studies be initiated as necessary to meet the scientific objectives of this strategy.*

4. The technical challenges intrinsic to detailed astronomical observation of evolving preplanetary systems and their precursor materials are

severe. They mandate use of the most sensitive existing observational facilities, improvement of current instrumentation, and development of more advanced telescopic imaging and spectroscopic detection systems. Utilization of the next generation of such observational systems—notably HST, SIRTF, SOFIA, and large ground-based optical and radio telescopes and arrays (and other facilities under development or planned by NASA, NSF, and other U.S. agencies, and by other nations)—will be critical for meeting the scientific objectives of the recommended program in this area. *COMPLEX therefore encourages the following multi-disciplinary activities between the responsible divisions at OSSA: participation of planetary scientists in the design and building of future observatories and facility instruments and in the allocation of observing time at existing observational facilities; joint support for multidisciplinary scientific initiatives; and joint development of instrumentation for extrasolar observation.* In consideration of the highly interdisciplinary nature of the recommended program and the anticipated costs of implementing it, *the committee further advises that relevant funding agencies, including separate programmatic divisions within NASA, should jointly support the cross-cutting technological developments and multidisciplinary scientific initiatives required for these extrasolar observations and studies.*

5. *The committee recommends pursuit of long-range instrumental and strategic initiatives that are conceptually applicable and potentially valuable to the investigation of extrasolar planetary materials in later stages of reconnaissance or in subsequent phases of exploration and intensive study, but that at present are technologically or theoretically too undeveloped to be of immediate utility in implementing the short-term strategy proposed in this report.*

Appendix A
The Large Star Sample Required for A Photometric Planetary Search

Consider a planet in a circular orbit, of radius a, around a main-sequence star of mass M_* and radius R_*. An empirical relation between stellar size and mass is given approximately by $R_* \sim M_*^{0.78}$. The amplitude of an occultation is just the ratio of the cross-sectional areas, or

$$A = 1.3\rho^{-2/3} M_P{}^{2/3} M_*{}^{-1.56},$$

where M_P and M_* are in solar units and ρ is the mean density of the planet, which could be somewhere in the range $0.5 < \rho < 8$ g cm^{-3}. A minimal detection of the occultation requires that $A \cong 3\sigma$, where σ is the photometric fractional uncertainty in the measurements. In the case of the Sun, this detection would require $\sigma = 3 \times 10^{-5}$ for an Earth occultation ($\rho = 5.5$, $M_P = 3 \times 10^{-6}$) and $\sigma = 3 \times 10^{-3}$ for a Jupiter occultation ($\rho = 1.3$, $M_P = 10^{-3}$). In the following, $\rho = 2.1$ is used, which yields the geometric mean value for $\sigma(\rho)$; the above range of values for ρ would yield values for σ within a factor of 2.6. The corresponding required photometric uncertainty is

$$\sigma \leq 0.26 M_P{}^{2/3} M_*{}^{-1.56}$$

If Poisson statistics apply, recording the event requires that $1/\sigma^2$ photoelectrons be counted. Assuming measurements in a 1000-Å visible passband with a 1-m telescope operating at 5 percent quantum efficiency, the on-target integration time required per star is

$$t = 3 \times 10^{-9}\sigma^{-2}10^{0.4m_v}\text{seconds},$$

where m_v is the apparent visual magnitude of the star. For main-sequence stars, the apparent magnitude can be reckoned from the mass and distance:

$$m_v \sim -8.4 \log M_* + 5 \log r + 1, \text{and so}$$

$$t = 10^{-7}r^2M_*{}^{-0.24}M_P{}^{-4/3}\text{seconds}.$$

For occultations of a 0.3-$M_\odot$ star by the Earth and Jupiter viewed from 10 parsecs, t = 300 s and 0.13 s, respectively.

The maximum duration of an occultation is 0.54 $a^{1/2}$ $M_*{}^{0.28}$ days, where a is in astronomical units, and it occurs once in an orbital period $P = a^{3/2}$ $M_*{}^{-1/2}$ yr. In order not to miss an event, each program star must be measured about four times during the *minimum* event duration sought. For estimation purposes, the limiting planet path is chosen to be offset by 0.5 stellar radii from the star's center. Then all the program stars must be measured in time:

$$T = 0.25(\sqrt{3/2})\ 0.54\ a_{\min}{}^{1/2}M_*^{0.28}\text{ days} = 10^4a_{\min}{}^{1/2}M_*{}^{0.28}\text{ seconds},$$

where a_{min} is the inner boundary of the search. For a_{min} = 1 AU and M_* = 0.3, T = 7.1 × 10^3 s or 2 h.

The occultation visibility zone decreases with increasing a. The probability of the planet passing within 0.5 stellar radius of the star's center, as seen by an observer in a random direction, is $R_*/2a$ or 2.33 × 10^{-3} $M_*{}^{0.78}$ a^{-1}. To be confident that N_s stars have been sampled for planetary occultations, that number times the inverse of the probability of visibility must be observed in the monitoring program, or

$$N_o = N_s \times 430a_{\max}M_*{}^{0.78},$$

where a_{max} is the outer boundary of the search. For a_{max} = 1 AU and M_* = 0.3, N_o = 1.1 × 10^3 N_s.

The survey program must therefore measure N_o stars, each for an integration time t, in a total cycle time T. Assuming an observational efficiency of 0.25, this requires

$$4N_ot < T.$$

For a typical star (M_* = 0.3) and a program duration of 10 yr (P < 10, a_{max} = 3.1) this requires

$$a_{\min} > 7.3 \times 10^{-10} M_P{}^{-8/3} r^2.$$

For the extreme case one takes $a_{\min} = a_{\max}$ and $M_P = M_P(\max) = 2 \times 10^{-3}$. For a minimum N_s of 10, this requires $r < 13$ parsecs and

$$N_o = 3.4 \times 10^3 N_s = 3.4 \times 10^4 \text{ stars}.$$

However, there are only about 1000 stars within 13 parsecs, so this method could not give an accurate estimate of the fraction of even Jupiter-size planets occurring around nearby stars.

Appendix B
List of Members
Commission on Physical Sciences, Mathematics, and Resources

Norman Hackerman, Robert A. Welch Foundation, *Chairman*
Robert C. Beardsley, Woods Hole Oceanographic Institution
B. Clark Burchfiel, Massachusetts Institute of Technology
George F. Carrier, Harvard University
Ralph J. Cicerone, University of California at Irvine
Herbert D. Doan, The Dow Chemical Company (retired)
Peter S. Eagleson, Massachusetts Institute of Technology
Dean E. Eastman, IBM T.J. Watson Research Center
Marye Anne Fox, University of Texas at Austin
Gerhart Friedlander, Brookhaven National Laboratory
Lawrence W. Funkhouser, Chevron Corporation (retired)
Phillip A. Griffiths, Duke University
Neal F. Lane, Rice University
Christopher F. McKee, University of California at Berkeley
Richard S. Nicholson, American Association for the Advancement of Science
Jack E. Oliver, Cornell University
Jeremiah P. Ostriker, Princeton University Observatory
Philip A. Palmer, E.I. du Pont de Nemours & Company
Frank L. Parker, Vanderbilt University
Denis J. Prager, MacArthur Foundation
David M. Raup, University of Chicago
Roy F. Schwitters, Superconducting Super Collider Laboratory
Larry L. Smarr, University of Illinois at Urbana-Champaign
Karl K. Turekian, Yale University